FACULTY DEVELOPMENT FOR VALID CLINICAL EVALUATION IN DENTAL HYGIENE EDUCATION

BY

JUDY SKELTON

A DISSERTATION PRESENTED TO THE GRADUATE COUNCIL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL
FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1983

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the cooperation, assistance, and support of many people. The writer herein expresses her deep and sincere gratitude to all of her academic advisors and friends.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS ii
LIST OF TABLES vi
LIST OF FIGURES vii
ABSTRACT viii

CHAPTER

I. INTRODUCTION 1
  Statement of the Problem 4
  Hypotheses 5
  Need for the Study 6
  Definitions 11
  Limitations 13
II. REVIEW OF THE LITERATURE 15
  Clinical Evaluation 15
  Instructional Design 26
    Principles of Faculty Development 26
    Characteristics of the Adult Learner 31
    Self-Instructional Design Guidelines 35
    Principles of Psychomotor Skill Learning 38
    Principles of Concept Attainment 40
  Summary 44
III. METHODOLOGY 46
  Experiment Design 46
  Instrumentation 47
    The Model 48
    Instructional Models 49

CHAPTER Page
Pilot Program 52
    Procedures 53
    Summary 57
Sample Population 57
Analysis of Data 58
Procedures 62
IV PRESENTATION OF FINDINGS 63
Introduction 63
Hypothesis 1 64
Hypothesis 2 67
Hypothesis 3 70
Survey Findings 77
Additional Findings 79
    Time Requirements 79
    Types of Errors 81
    Professional Backgrounds of Faculty 83
Summary 84
V DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS 85
Introduction 85
Discussion and Conclusions 87
    Hypothesis 1 and 2 87
    Hypothesis 3 90
    Survey Findings 91
    Additional Findings 93
    Summary of Conclusions 95
Recommendations for Further Study 95
    Dentition Models 96
    Training 96
    Statistical Analyses 97
    Additional Findings 97
Closing Statement 98

APPENDICES

A ADMINISTRATIVE AND PARTICIPANT INSTRUCTIONS FOR CALCULUS DETECTION MODULE 100
B SURVEY 109
C INSRUCTIONS FOR DENTION MODEL CONSTRUCTION 113

Page
REFERENCES 118
BIOGRAPHICAL SKETCH 124

LIST OF TABLES

Table		Page
1	Summary of Research on Training to Improve Clinical Evaluation	19
2	Pilot Program Results Correct Responses and Individual Percentage of Agreement with Criteria	54
3	Time Requirements for Pilot Program Pretests and Posttests	56
4	Pretest Scores of Individual Faculty for Calculus Detection Reported as Percentage of Agreement	66
5	Classification Matrix for Cohen's Kappa . .	71
6	Posttest Scores of Individual Faculty for Calculus Detection Reported as Percentage of Agreement	73
7	Comparison of Individual Pretest and Posttest Scores	74
8	McNemar Data Summary Matrix	75
9	Summary of Data for McNemar Statistic . . .	76
10	Averaged Time Requirements for Pretests 1 and 2 and Posttests 1 and 2 for Individuals and as a Group	80
11	Summary of Errors	82

LIST OF FIGURES

Figure — Page

1 Relations between phases of learning and events of instruction 27

2 Knowledge bases used to develop calculus detection training program 29

Abstract of Dissertation Presented to the Graduate Council
of the University of Florida in Partial Fulfillment
of the Requirements for the Degree of Doctor of Philosophy

FACULTY DEVELOPMENT FOR VALID CLINICAL EVALUATION IN DENTAL HYGIENE EDUCTION

By

Judy Skelton

December 1983

Chairperson: Dr. Albert B. Smith
Major Department: Instructional Leadership and Support

Validity of clinical evaluation in dentistry is minimal. The purpose of this study was (1) to determine the level of validity in one dental skill, calculus detection, and (2) to determine whether one type of systematic training could improve evaluative judgment of the faculty involved. Three hypotheses were proposed, postulating that the calculus detection skill levels of dental hygiene faculty were low by individual and by group, and that training would improve evaluation skills and thus, the validity of judgment in calculus detection.

The sample population studied was 21 instructors from three dental hygiene programs (Pensacola Community College in Pensacola, Florida; Tallahassee Community College in Tallahassee, Florida; and Raymond Walters College in Cincinnati, Ohio). The training program for faculty was a self-instructional module consisting of simulated periodontally involved dentition models and a series of self-graded practice activities.

A designated project administrator at each school administered two pretests for ascertaining initial calculus detection skill levels. The faculty participants were then instructed to practice until each could repeatedly obtain an accuracy level of at least 90%. When 90% accuracy had been achieved, the participants were to request the posttests.

The investigator analyzed results of the pretests using percentage of agreement and Cohen's Kappa measurements to determine levels of accuracy. McNemar's test for significance of change was used to compute significance in changes of performance. From the data obtained, it was statistically determined that (a) only one of 21 faculty could detect calculus at 90% accuracy on the pretests, (b) dental hygiene faculty as a group were unable to accurately identify calculus at a 90% level of accuracy, and (c) dental hygiene faculty improved significantly in the ability to detect calculus from pretest to posttest.

Additional research is required to identify the relationship between training and validity in clinical evaluation in dental hygiene. Although training appeared to have an effect on calculus detection skills, it was difficult to ascertain if the demonstrated increase in skill reflected real differences or limitations of the study design.

CHAPTER I

INTRODUCTION

Clinical evaluation is a unique and difficult task of "collecting information on student clinical performance in order to make informed decisions regarding student progress . . . ." (Morgan and Irby, 1978, p. xi). More precise and objective clinical evaluation is needed in all health-related education programs. The purpose of this research was to examine one particular field--clinical evaluation in dental hygiene education.

Several investigators have demonstrated a clear need for more research in clinical evaluation techniques in dentistry (Abou-Rass, 1973, Houpt and Kress, 1973; MacKenzie, 1974, O'Connor and Lorey, 1978; Patridge and Mast, 1978). The task is complex because of the myriad of clinical competencies required in dentistry and the large number of variables that influence evaluation.

Clinical evaluation of psychomotor skills is a primary task in dental hygiene education programs. Students perform patient services which require fine motor activities; the results of students' efforts must be judged adequate or inadequate by dental hygiene instructors. Evaluation to ascertain clinical competency in dental hygiene skills is a complex task with no simple or prefabricated system that be applied to all settings because of the many variables in an evaluation system and since many of the skills evaluated are often a combination of motor activities and judgment abilities that cannot be positively confirmed at the time the task is completed.

Choosing this topic for study was an effort to determine some of the complexities involved in clinical evaluation. The skill of interest in this investigation was detection of deposits of calculus on the crowns and roots of teeth. Because the deposits are often beneath the gingival tissues, detection must be accomplished without vision through the use of a dental exploring instrument designed to be an extension of the therapist's fingertips (Pattison and Pattison, 1979; Schulger, Yuodelis and Page, 1977; Wilkins, 1976; Woodall, Dafoe, Young, Weed-Fonner, and Yankell, 1980). A deposit-free subgingival tooth surface can be verified only by extraction or surgical procedures, and subsequent visual observation which is antithetical to the purposes of dental hygiene therapy.

No ideal solution to this dilemma of evaluation without positive confirmation exists, therefore, the tactile discrimination skills of faculty are assumed to be valid and reliable to the point that decisions about a student's competency could be based upon unverifiable skills of the faculty. The purpose of this study was to ascertain the level of validity in calculus detection judgment of dental hygiene instructors using dentition models and to attempt to improve faculty skills through training. In this study, validity was defined as a measure of agreement with a criterion (see Definitions, p. 11); the criterion was the absence or presence of calculus on dentition models. Reliability was defined as consistency of decision-making across repeated measures of parallel test forms. This similarity and relationship of validity and reliability is intrinsic to the research questions and statistical measurement employed in the study; thus, validity is the descriptive term used in the text.

The design of this research attempted to control two variables in clinical evaluation of calculus detection--patient differences and training method--and to identify more accurate statistical measurements for the problem. Many other variables need to be investigated such as time allowed for evaluation, different methods of training, and choice of instruments.

The theoretical framework for this investigation was a compilation of several knowledge bases. Gagne's model for identifying relationships between the phases of learning and the events of instruction (Gagne and Briggs, 1974, p. 285) was the foundation. Elements of instructional theory for teaching psychomotor skills were combined with theoretical models for concept attainment and self-instruction. Additionally, the tenets for adult learning and faculty development were applied in the design and implementation of the instructional component.

Statement of the Problem

According to the literature available, validity is minimal and reliability is low in clinical evaluation in dentistry (Abou-Rass, 1973, Saratti, 1974; Hiller and Kerber, 1980; Fuller, 1972; Gaines, Bruggers, and Rasmussen, 1974; Hinkelman and Lang, 1973; Houpt and Kress, 1973; Hunter, Salkin, Leve, and Hildebrand, 1975; MacKenzie, 1973; O'Connor and Lorey, 1978; Partridge and Mast, 1978). At least three factors-- poor scoring systems, unclear performance criteria and inadequate training--have been identified as contributing to minimal validity and low reliability in clinical evaluation (Patridge and Mast, 1978). The focus of this study was on the effect training had on validity (and reliability) in evaluation.

The specific problem of this investigation was, first, to determine the level of validity in calculus detection by dental hygiene instructors and, second, to determine whether one type of systematic training could improve the validity of instructor evaluations. As stated previously, validity was defined as a measure of agreement with known (valid) criterion; in this study the question was, can faculty accurately and consistently (reliability) find calculus that is known to be present (validity).

Clinical evaluation in dental hygiene involves detection of deposits through the use of an instrument on tooth surfaces beneath the gingival tissues. Therefore, in order to achieve validity in clinical evaluation, evaluators must be accurate in detection of subgingival deposits on tooth surfaces. This inference led to the objectives of this study. The first objective was to ascertain the level of validity in the calculus detection skills of instructors in dental hygiene programs through a criterion-related pretest measure The problem was to identify the extent to which dental hygiene instructors could judge whether calculus was present.

Further objectives of this study stemmed from the investigator's desire to improve clinical evaluation in dental hygiene The second objective was to develop a training program for faculties which would improve their calculus detection skills and thus their ability to do clinical evaluation. Subobjectives to this second goal were (a) to identify the educational components of calculus detection and (b) to combine the current knowledge in those areas into a successful training program.

Hypotheses

From these objectives, hypotheses were developed. Testing of the hypotheses was accomplished by using the null form and statistically measuring for levels of validity before and after training.

H_o1: Dental hygiene instructors as individuals are able to accurately identify calculus when criteria are available for comparison.

H_o2: Dental hygiene instructors as a group are able to accurately identify calculus when criteria are available for comparison.

H_o3: Training will not significantly improve the ability of dental hygiene instructors to detect calculus.

Definitions and statistics are explained in later sections.

Need for the Study

The needs for this study fall into two categories: (a) the task-specific needs of improving calculus detection skills, and (b) the broader needs related to improving clinical evaluation.

Use of an exploring instrument to differentiate between pathologic and non-pathologic tooth surface irregularities is one of the most strategic skills a dental hygienist must learn. The skill involves tactile discrimination abilities which require tactile sensations of deep pressure and spatial localization elements transmitted by the nerve fibers that terminate just under the surface of the skin

(Troy, 1971). This complex psychomotor skill must be refined to the point that an operator can discriminate between the many surface variations such as the cemento-enamel junction, root bifurcations and other concavities and convexities found subgingivally. Proper dental therapy cannot be delivered without a highly developed ability to discriminate tactilely (Pattison and Pattison, 1978; Wilkins, 1976; Woodall et al., 1980).

Calculus detection through tactile discrimination is difficult to teach and to learn. Most dental texts and curricula locate instruction for calculus detection in the initial phases of the program. Faculty arrange formal practice sessions to facilitate skill acquisition. However, the standards of criteria for the skills may be unclear and instruction suboptimal because faculty have a low percentage of agreement in detecting calculus. The investigator has also noted that formal training after dental hygiene school is rare. The significance of that observation for this study is that faculty of dental programs do not receive systematic training or reinforcement of their skills; yet feedback and reinforcement are basic principles of learning and skill maintenance (Ausubel, 1968; Davies, Alexander, and Yelon, 1974; Gagne, 1974). Therefore, if a systematic method for developing, maintaining and reinforcing calculus detection skills can be developed, both students and faculty will benefit.

From a broader perspective, the methods employed in this research to improve faculty skills should be effective and could be applied to other areas.

The second category of needs mentioned previously is improvement of clinical evaluations. Although this category is broad, some essential needs in clinical evaluation apply to this study. The first is the issue of the quality of health care. If a student cannot be accurately judged as competent, the standards of care may be compromised. One responsibility of the instructional faculty in health care education programs is to ensure competence.

The second need related to clinical evaluation involves the rudimentary purposes of evaluation in general. The purpose of evaluation is to make decisions about the competence of students and therefore about their progress toward their future goals. These decisions must be based on sound, valid information. If the clinical evaluation system is not valid, the decisions based on the results of the system are arrived at inefficiently or, worse, are unfair to certain individuals and to society.

A third need for improving clinical evaluation involves practicality. Knowledge in dentistry as in all other areas of health is increasing rapidly. A strongly valid system of evaluation can increase the efficiency in learning in three ways. The student who is accurately judged to be competent will perform future tasks based on learned skills at the required standard of quality without having to go back and relearn. Second, through an evaluation system, instructors can increase efficiency in learning by accurately identifying where incompetency exists and channelling attention toward a specific area. Third, errors in measurement require more observations; therefore, if the evaluation is accurate, fewer measurements will be required and efficiency will be increased.

To date, the methods employed to improve clinical evaluation (reviewed in the literature in a subsequent section) have failed to control adequately for the many variables affecting validity and reliability. Petridge and Mast (1978) identified three general factors: poor scoring systems, unclear performance criteria and inadequate training. Within each of these broad areas many variables need to be investigated. Training of faculty was the main variable of interest in this study; issues addressed included methods of training and measurement. The literature also revealed that little attention has been given to the principles of faculty development, adult learning and instructional design.

In addition, previously used statistical methods for measuring validity and reliability in clinical evaluation may have resulted in spurious results. This investigator focused on the ability of faculty to match criteria (i.e., it was known whether calculus was present or not). As faculty judgment approached agreement with the criteria, evaluations became more valid. Under certain circumstances faculty evaluations could become highly reliable without being valid. For instance, instructors might agree that the sensation they felt from root roughness was calculus when in reality it was not. The agreement would be high, but the validity for detecting calculus would be low. Another example of the weaknesses in statistical methods used previously is the lack of control for chance and rare-events. Given only two choices there could be a 50% agreement by chance and if calculus is present on only 5% of the surfaces available, agreement could occur 95% of the time without the subject's giving a response. The measuring procedures used in this study circumvented this problem of spuriously high intrarater and interrater agreement.

In order to evaluate the clinical skills of learners, faculty must possess valid and reliable evaluative skills. Validity is rare in dental clinical evaluation (Abou-Rass, 1973, Biller and Kerber, 1980; Fuller, 1972; Gaines et al., 1974; Hinkleman and Lang, 1973, Houpt and Kress, 1973; Hunter et al., 1975; MacKenzie, 1973; MacKenzie, 1974; O'Connor and Lorey, 1978; Petridge and Mast, 1978) and educators need a means of improving validity. This researcher attempted to meet the need by developing a way to achieve greater validity in faculty performance through training

Definitions

Self-Instructional Training Module is a self-contained unit of materials which provides most of the instruction allowing for variation in time (self-pacing) and individual skill levels. Materials in the module for improving calculus detection skills includes brief written introductory statements, instruction on how to proceed through the practice exercises, a physical model of the adult dentition and check sheets. (For samples of these items, see Appendix A).

Psychomotor Skill is a coordinated series of muscular movements in response to sensory perception (synonymous with sensory-motor, perceptual motor). The psychomotor skill of interest in this investigation was calculus detection; a finely tuned skill involving tactile discrimination through a metal instrument held in the hand of the operator

Model is a life size physical representation of the human dentition and surrounding soft tissues.

Calculus Detection is the identification of a hard mineralized deposit (calculus) on the surface of a tooth through the use of an explorer (dental instrument).

Dental Hygiene Faculty are instructors from three different dental hygiene programs.

Clinical Evaluation is a "collection of information on student clinical performance in order to make informed decisions regarding student progress" (Morgan and Irby, 1978, p. xi). In this study, the information collected will be the number of correct and incorrect responses when subjects attempt to identify the location of calculus deposits on simulated dentition models.

Validity is the extent to which an instrument measures what it is intended to measure (Mehrens and Lehman, 1973). Validity of a judgment is the extent to which the judgment is consistent with the actual situation. In this study validity was specifically defined as a measure of agreement with known criteria--the presence or absence of calculus. According to Ary, Jacobs, and Razavieh (1979), this is criterion-related validity; it is the relationship between the scores on a measurement instrument and a variable believed to directly measure the characteristic in question.

Reliability is the degree of consistency of a measuring instrument. In this study reliability was further defined as consistency of decision making across repeated measures of parallel test forms.

Accuracy for this study was defined in two ways. First, it was the percentage of agreement with the criteria set to be .90 (90%). Second, accuracy was the resulting coefficient as measured by Cohen's (1960) Kappa.

Criterion is the actual presence (or absence) of calculus on the teeth of a dentition model. The calculus can be verified by record of placement, by sight and, if necessary, by removal.

Limitations

The experimental format employed for this research was a quasi-experimental (non-randomized) pretest-training-posttest design (Ary et al., 1979). Selection of this design was predicated on the lack of resources (financial and time) to obtain a control group; the faculty available to participate did not live near the researcher or each other. Inability to randomly select subjects restricts the generalizability of the results to all dental hygiene instructors.

In addition, the faculty participants were volunteers. The investigator recognized the influence of further restricting the possible inferences. However, this compromise was unavoidable and was somewhat desirable in regard to the theories of adult education and faculty development. Adults want to be self-directed, and for faculty development programs to be successful, faculty must perceive a need to change.

A third limiting factor involved the use of the physical model of the mouth for refining a psychomotor skill instead of using a human mouth. The nature of the task of calculus detection is complex in that either the true results are not known with any assurance, or a model must be used to confirm the presence or absence of deposits. The researcher chose to use a model and made every attempt to create one which closely resembles the oral structures, contours, and textures. Transfer of learning has been shown to be successful when the simulations used for training closely resemble normal conditions (Hall, Goldberg and Steen, 19[illegible]).

CHAPTER II

REVIEW OF THE LITERATURE

Clinical Evaluation

An extensive review of the literature on clinical evaluation in dental education confirmed the need for further research into the area. All of the studies reviewed advocated continued study into methods for improving dental clinical evaluation, citing critical deficiencies in inter-rater and intrarater reliability, validity, rating scales, and practicality (Abou-Rass, 1973; Biller and Kerber, 1980; Gaines et al., 1974; Hinkleman and Lang, 1973; Houpt and Kress, 1973; Hunter et al., 1975; MacKenzie, 1973; MacKenzie, 1974; Morgan and Irby, 1978, Natkin and Guild, 1967; O'Connor and Lorey, 1978; Rasmoen, Shotwell, and Lang, 1978).

Considerable agreement was also found about why improvement in clinical evaluation is necessary. The following list is a composite of the reasons found in the literature (Abou-Rass, 1973; Baratti, 1974; Hinkleman and Lang, 1973; MacKenzie, 1973; Natkin and Guild, 1967; Ormes, 1973):

1. To maintain the quality of health care in general. If competence cannot be accurately judged, then there is an increased possibility that performance standards

will fail. For example, if a student does not receive consistent feedback about the calculus deposit remaining on the distal surface of maxillary posterior molars, that person may leave calculus in the future. This lack of competence contributes to the decline of the periodontal health of all patients treated by the practitioner.

2. To increase efficiency. As dental knowledge increases, so do the procedures which students must master. A valid evaluation system will maximize proficiency of skill acquisition thus, minimizing additional performance demands. If instructors do not always agree on whether root roughness is acceptable or unacceptable, more trials must take place before consensus can be reached about the student's competence. "When the consequence of behavior is unreliable and inconsistent, learning tends to be unpredictable and inefficient" (Natkin and Guild, 1967, p. 153). Feedback to students is the single most important variable that governs skill learning (Winkleman and Lang, 1973).
3. To achieve fairness in decision making about students' abilities. Evaluation systems have as their purpose identification of directional alternatives for students; if the clinical evaluation system used is not adequate, valid and reliable decisions cannot be made about the progress of students on individual tasks and through the curriculum

4. To minimize students' confusion, frustration and fear. Applying behavioristic principles, inconsistent grading (consequence after a behavior) is likely to produce undesirable emotional conditions in the learner; likewise, a reduction in the subjectivity of grading will lead to more confidence and responsibility in the student. "One of the most reliable methods of producing experimental neurosis in lab animals is to confront the animals with a problem situation in which no consistent consequence follows attempted solutions" (Netkin and Gelld, 1967, p. 153).

Having reconfirmed the need for study, the focus of the literature review (on clinical evaluation) moved toward identifying similar studies that supported or refuted the hypothesis that training can increase the skills of faculty. A meager amount of literature was available that directly addressed training. Most studies did not focus on faculty training as a way to improve clinical evaluation or defined faculty training as training in the use of the evaluation instrument, or both.

Petridge and Mast (1978) reviewed the research on dental clinical evaluation and developed a conceptual framework that is helpful in analyzing the subject. They first defined clinical evaluation as a type of work-sample test which "provides a controlled trial of the examinee's performance under specific conditions" (p. 300). Further definition was provided by Ausubel (1968); according to him, work-sample tests assess competencies in direct and therefore valid methods.

The second objective accomplished by Patridge and Mast was the categorization of independent variables in experimental clinical evaluation studies. These categories were identified as being associated with work-sample tests—performance criteria, scoring systems and training of evaluators.

Training of evaluators was the independent variable of most concern to this researcher. In reference to work-sample tests, both interrater and intrarater validity increases when individual bias is minimized through training of evaluators; therefore, training is considered essential (Fitzpatrick and Morrison, 1971).

This claim was debated in the dental literature. Results were inconsistent in studies which included reference to training and its effect on clinical evaluation. Of the 18 reports on clinical evaluation reviewed, 11 made reference to training (see Table 1), and each possessed different characteristics.

Several different types of training were described. The types ranged from one 15-minute session (Gaines et al., 1974) to a rather extensive series of eight discussion sessions (Natkin and Guild, 1967). All of those described indicated the format for training was primarily a discussion with some formats using practice exercises in addition to the discussion. The focus of the discussion included using

Table 1

Summary of Research on Training to Improve Clinical Evaluation

Author(s)	Research Question(s)	Methods	Training	Results
Abou-Rass 1973	More reliable clinical evaluation instruction can be developed using a detailed task analysis of endodontic therapy	-4 raters rated 24 endodontic cases -the same 4 raters rated 34 cases after training	5 sessions over an 8 week period with emphasis placed on definition, measurement	Overall agreement increased from 54% to 77%. Concluded that both overall agreement and agreement on individual criteria increased as a result of training
Benatti 1974	Homogeneity of grading practices would increase after training	-8 dental hygiene faculty -15 patients of different levels of difficulty -faculty groups graded cases after students had assessed patient's mouth	3 training sessions: a review examination techniques b practice detection on models with varying roughness c practice detection on patients to maximize transfer of learning	No consistent pattern of grading homogeneity occurred. Concludes training sessions ineffective
Fuller 1972	Consistency in the evaluation of a Class II amalgam preparation or Class III goldfoil preparation could be improved with the use of a model with a checklist during evaluation	-4 male raters -3 experimental conditions A traditional glances and grade B model with checklist of training 25 lab practical products of amalgam and gold cavity preparations were evaluated	2-hour discussion sessions on concepts and techniques taught to students; discussion of human factors affecting student's work; an explanation of the rating techniques and practice in rating	There was no significant increase in interrater reliability when using training with a model and checklist

Continued

Table 1

Continued

Author(s)	Research Question(s)	Methods	Training	Results
Gaines, Bruggers, Rasmussen, 1974	The type of rating scale will influence reliability.	7 raters evaluated 8 wax carvings using a subjective rating scale (outstanding average poor) 3 months later same 7 evaluators rated different wax-ups using an objective scale (completely smooth slightly rough major nicks and scratches)	Second session had 15-minute training session.	Analysis of variance showed a significant difference between raters when using the subjective scale but not when using the objective scale (Authors speculated that the training helped and therefore length of training was not important)
Goepford and Kerber, 1980	Comparison of 2 methods for evaluating primary Class II cavity preparation	-15 raters (dentists) evaluated 24 Class II cavity preparations in ivorine teeth using 2 scales: 1 5-point scale A textbook quality B minor corrections needed C major corrections needed D substantial faulty input E irreversible damage	Training introduced after 2 pieces-and-grade sessions. Consisted of an explanation of the new analytical grading system. Presentation included slides on exemplars/non exemplars.	Pearson's product moment correlation used to establish intrarater reliability ANOVA used for inter-rater reliability Both reliabilities increased but the author felt the results were spurious due to repeated evaluations of the same preparations. No mention made of training effect

Continued

Table 1

Continued

Author(s)	Research question(s)	[illegible]	Training	Results
		2. Glance-and grade 16 criteria each preparation graded 4 times by all raters in random order		
Hinkelman and Long 1973	Comparison of the reliability of a 3-point scale with a 2-point scale for evaluating inlay preparations	-4 evaluators, 68 inlay preparations randomly selected from 124 students -2-point scale pass/fail -3-point scale a no improvement necessary b minimally acceptable c clinically unacceptable and uncorrectable	Training provided but amount and type not described in detail. Report referred to the fact that the evaluation instrument was used throughout a school term to familiarize faculty with it and that during a week long break between the evaluation of the 1st and 2nd group of 34 preps, a training session was held to reinforce criteria	More agreement was found in the pass/fail system indicated by an increase in percentage of agreement measures. Concluded training helped slightly and recommended frequent recurrent training
Houpt and Kress 1973	Comparison of the reliability of a 2-point scale and a 5-point scale (with refined points) for cavity preparations in dentals	-3 instructors, 36 dental students, 16 dental assistants divided into 3 mixed groups -2 rating scales a. 2-point scale b. 5-point scale with partial definitions	3 training sessions a. 5 specimens rated no feedback b. 5 specimens with overall feedback on 2 and individual criteria feedback on 3 c. 2 specimens regraded correct responses and error descriptions provided	Reliability of total product good, however, specific criteria ratings had low reliability. There was no significant effect of training on reliability

Continued

Table 1

Continued

Author(s)	Research Question(s)	Methods	Training	Results
		c. 5-point scale with all points defined 8 criteria for cavity preparation -5 teeth rated on 8 criteria totaling 40 judgments -ratings obtained repeatedly for 3 sessions		
Natkin and Guild, 1973	Reliability of preclinical laboratory evaluation in endodontics can be improved	5 raters (4 graduate students, 1 full-time faculty) 3 attempts (before exposure to evaluation system, after reading the system and after training sessions) lab samples of 5 categories of endodontic procedures, 4 rating categories: a. slight b. moderate c. serious d. critical	4 discussion sessions in which 10 randomly-selected student endodontic projects were reviewed and evaluated jointly	Statistically significant increase in reliability (measured by the mean of all the average durations in the study) only achieved after the raters were provided with a formal evaluation system; no significant increase reported after training

Continued

Table 1

Continued

Author(s)	Research Question(s)	Methods	Trials	Results
O'Connor and Lorey, 1978	The use of comparison stimuli can improve interrater agreement (two studies).	I. 16 faculty evaluated 40 specimens of molar chamfer preparations in melamine teeth 6 criteria 5-point scale -faculty randomly divided into 2 groups (A and B). A evaluated teeth B evaluated using slides as exemplars.		ANOVA reliability estimate revealed directionally reduced variances using exemplars on 2 criteria and total score.
		II 18 judges evaluated 20 pin-ledge preparations 7 criteria 5 point scale -judges divided into 2 groups (C and D). C, evaluated with exemplars D evaluated without exemplars		Study II failed to support study I; exemplar effectiveness not confirmed.

an evaluation form, clarification of performance criteria and utilization of a scoring system.

The Baratti (1974) study was the only effort found to train faculty in using the motor skills and tactile discriminations required for calculus detection. Training sessions in this investigation were a series of three practice sessions, each offering different activities. As reported in Table I, Baratti found no improvement in the evaluation procedures. Although a variety of instructional techniques were reported, none of the studies reviewed indicated the use of a self-instructional method. Evidence in the dental literature indicates that this technique is being successfully integrated into dental curricula for teaching students (Williams, 1981), but this researcher did not find any evidence of self-instructional techniques being used in faculty training for increased reliability in evaluation.

In addition to the differing instructional techniques described, a great variety in the products being evaluated was evident. Many products were models of some type of restorative tooth preparation in ivorine teeth and the evaluative skills required were primarily judgment skills based on visual cues. A dental instrument may have been used during the examination, but apparently visual confirmation was always available as the primary sensory mechanism. Again, the Baratti study (1974) was the only research found which involved the topic of calculus detection or any other subject paralleling the evaluative skills being investigated.

Because of the differing features of the available research, comparisons and conclusions are difficult to identify, and generalizations are impossible. Some studies report that training had no effect (Baratti, 1974; Fuller, 1972; Houpt and Kress, 1973, Natkin and Guild, 1967; O'Connor and Loray, 1978). Others report that training improved clinical evaluation (Abou-Rass, 1973) and still others speculated training had a positive affect (Gaines et al., 1974; Morgan and Irby, 1978). No consistant pattern emerged.

Few inferences can be made from the available literature to aid in answering the questions of this study. Of the research reviewed, only one study dealt with calculus detection and the skills required, training was rarely investigated as a primary variable, and the training referred to in the literature was not defined as skill development but as calibration on the use of an evaluation instrument. The effectiveness of training can vary greatly depending on how instructional variables are employed. From general experience the inference can be made that well designed training should increase skills in clinical evaluation.

A review of the literature on clinical evaluation revealed a weakness in the training programs employed to improve clinical evaluation; therefore, a literature search was conducted to find an acceptable instructional approach for improving faculty calculus detection skill. The most appropriate model found was developed by Gagne (1977) and identified the relationships between the phases of learning and the events of instruction. This model was used as a guideline for developing the instructional sequence for the calculus detection training unit (see Figure 1). The column labeled Corresponding Activities identified the instructional events developed for the training program in this study. Detailed information on the module is provided in Chapter III and Appendix A.

Instructional Design

The activities which were selected for use in the training program were developed from an eclectic theoretical framework. Figure 2 illustrates the knowledge bases included.

Review of the literature on these five subjects was limited to recent publications. The objectives for the research process were to identify current knowledge and practice. No attempt was made to review all of the literature on each of the topics.

Principles of Faculty Development

Review of the literature on clinical evaluation reconfirmed the need for more research in the area, revealing a recognition of inaccuracy of evaluation. One

Learning Phase	Instructional Event	Corresponding Activities
Motivation Phase EXPECTANCY	1. Activating motivation 2. Informing learner of objectives	1. Small group discussion 2. Pretest
Apprehending Phase ATTENTION: SELECTIVE PERCEPTION	3. Directing attention	3. Distribution of module packages containing modules, instruments, etc.
Acquisition Phase CODING; STORAGE ENTRY	4. Stimulating recall 5. Providing learning guidance	4. Provided written instruction 5. Provided recording [illegible]
Retention Phase MEMORY STORAGE		

Continued

Learning Phase	Instructional Event	Corresponding Activities
Recall Phase RETRIEVAL	6 Enhancing retention	6. Repetition of practice sequences
Generalization Phase TRANSFER	7. Promoting transfer of learning	7. Practice sets designed to provide variety of experiences 8. Models constructed to simulate clinical conditions 9. Series of practice exercises with feedback
Performance Phase RESPONDING	8. Eliciting Providing feedback	10. Posttest
Feedback Phase REINFORCEMENT		

Note. Column 1 and 2 are events defined by Gagne (1977, p. 285), column 3 is the corresponding activities of the training program for this research.

Figure 1. Relations between phases of learning and events of instruction.

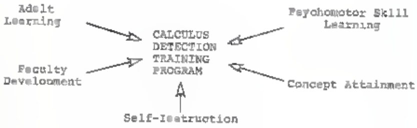

Adult Learning
Faculty Development
CALCULUS DETECTION TRAINING PROGRAM
Psychomotor Skill Learning
Concept Attainment
Self-Instruction

Figure 2. Knowledge bases used to develop calculus detection training program.

aspect of the approach taken in this study to improve the deficiency relies on the application of the principles of faculty development; the literature was surveyed to define and delineate the need for faculty development and to identify the currently known principles in the field.

The concept of faculty development is in a developmental phase (Francis, 1975; Gaff, 1975; Phillips, 1976), but for the purpose of clarity in this paper, the construct to be applied will be the same as that used by Smith (1976). He has proposed a three-pronged conceptual model based on Gaff's theories about approaches to faculty development (Gaff, 1975), the components of which are instructional development, personal development, and organizational development. This research employs the personal development approach, which focuses on the faculty member's growth by helping in the acquisition of new skills or refining existing ones with the ultimate goal of improving teaching efficiency (doing things better), and effectiveness (doing the right things better) (Hammons, Smith, and Watts, 1978).

Rationale for faculty development stems from the preceding definition; a need for increased efficiency and effectiveness occurs for the following reasons: declining student enrollments, dwindling tax dollars, and an increasing number of tenured faculty is reducing the number of new faculty positions, thus effecting faculty turnover. This condition predicates the need for development of existing faculty. The quality and success of educational programs depend in part on the ability of faculty to remain vital, to grow, to adapt to a new clientele of student, and to acquire and teach using new knowledge and technology (Centra, 1976; Gaff, 1975, Hammons et al., 1978). Additional purposes, identified by O'Banion (1977), include enhanced self-confidence and self-worth, greater competence in instruction in general, and increased competence in specific professional subject areas as reasons for faculty development programs.

Balanced against all the positive reports supporting faculty are an equal number of negative points. Many staff development programs occupy a low status as reported by Wood and Thompson (1980). The major reasons for this low status are as follows:

1. Educators report that many programs are poorly planned and organized, irrelevant, impersonal, and lack participatory involvement in planning and implementation. Also cited were lack of clear objectives and follow-up activities.

2. Many administrators take demeaning views of faculty as disliking inservice education, trying to avoid professional growth, requiring rewards to obtain cooperation, and preferring to be directed rather than self-directing. These views often become self-prophesying.

3. Many programs are too broad in scope, spreading across too many disciplines.

4. Knowledge about adult learners and how they learn is ignored in terms of the design of inservice programs.

5. No modeling is applied; inservice seminars are not designed around the principles which they advocate. (p. 375)

Current literature on faculty development offers design formats, which attend to the reported deficiencies listed in the preceding paragraph. Much of the literature reviewed includes summaries of available literature and was selected for inclusion because of relevancy.

Characteristics of the Adult Learner

First to be considered are the characteristics of the adult learner, which should direct the actual design of a program. Approaches to the adult as a student include the identification of different stages of adult development and the formulation of general principles, which describe how learning occurs (Knowles, 1978).

Hodkinson (in Gaff, 1975) details a three-stage transitionary profile of the adult, beginning with the younger years of 22 through 29. During this stage, the adult explores and seeks information about academic roles to be entered. In the second stage, career commitments are re-affirmed or a big shift occurs. The information sought relates to the chosen avenue of activity. In the later stage, the adult seeks order and stability with interest focusing on specific perceived deficiencies.

Characteristics of the adult learner described by Hammons et al. (1978) can be linked to these stages:

1. Adults learn what they want to learn; forcing them into rigid programs does not work. They must recognize a deficiency to motivate them to learn and to change

2. Adults become increasingly self-directed in their approach to learning, they want to assess their own goals and implement their own strategies.

3. Adults tend to focus on solving specific problems facing them in everyday work settings. (p. 120)

Complementing these characteristics is a summary of the literature on the adult learner by Wood and Thompson (1980)

> Adults will commit to learning something when the goal and objectives of the inservice are considered realistic and important to the learner, that is, job related and perceived as being immediately useful.
>
> Adults will learn, retain, and use what they perceive as relevant to their personal and professional needs.
>
> Adult learners need to see the results of their efforts and have accurate feedback about progress and their goals.
>
> Adult learning is ego-involved. Learning a new skill, technique or concept may promote a positive or negative view of self. There is always fear of external judgment that we adults are less than adequate, which produces anxiety during new learning situations such as those presented in inservice training programs.
>
> Adults come to any learning experience (inservice) with a wide range of previous experiences,

> knowledge, skills, self-direction, interests, and competence. Individualization, therefore, is appropriate for adults as well as children.
>
> Adults want to be the origins of their own learning, that is, involved in selection of objectives, content, activities, and assessment in inservice education.
>
> Adults will resist learning situations which they believe are an attack on their competence, thus the resistance to imposed inservice topics and activities.
>
> Closely related, adults reject prescriptions by others for their learning, especially when what is prescribed is viewed as an attack on what they are presently doing. Doesn't that sound like current inservice practice? We typically use inservice training to eliminate weakness we see in our personnel.
>
> Adult motivation for learning and doing one's job has two levels. One is to practice and do an adequate job. The second level is to become deeply involved, going beyond the minimum or norm. The first level of motivation comes as the result of good salary, fringe benefits, and fair treatment. The second builds on the first, but comes from recognition, achievement and increased responsibility--the result of our behavior and not more dollars.
>
> Motivation is produced by the learner; all one can do is encourage and create conditions which will nurture what already exists in the adult.
>
> Adult learning is enhanced by behaviors and inservice that demonstrate respect, trust, and concern for the learner (p. 376)

The characteristics of the adult learner should direct the design of faculty development programs. Several authors combine instructional principles with adult characteristics that result in recommendations for a program format. One group of researchers has applied Piaget's developmental stages model to John Dewey's experimental philosophy

(Keeton, 1976; McKinnan, 1976; Wood and Neil, 1978) The results led to the conclusion that the most efficient design for inservice education should include direct and concrete experiences where the learner applies what is being learned.

In the same article, the authors identify research that supports a belief that the climate of the learning situation is important to adults. Two differing research projects revealed that adults learn best in informal settings, which allows for social interactions among the learners (McKinnan, 1976, Rapport and Rapport, 1975).

To maximize the affect of training for increased reliability in evaluation, the instructional design employed must have the following characteristics:

1. The faculty must recognize the need for training.
2. The setting for the training should be informal and flexible to fit into existing schedules.
3. The instructional materials should provide for individualization in both learning style and pace.

In summary, what educators say about staff development and what is known about adult learners suggest that inservice educators should

1. include more participant control over the what and how of learning
2. focus on job-related tasks that the participants consider real and important
3. provide choices and alternatives that accommodate the differences among participants
4. include opportunities for participants in inservice training to practice what they

> learn in simulated and real work settings as part of their training
>
> 5. encourage the learners to work in small groups and to learn from each other and
>
> 6. reduce the use and threat of external judgements from one's superior by allowing peer-participants to give each other feedback concerning performance and areas of needed improvement. (Wood and Thompson, 1980, p. 377)

The information revealed by this limited literature review on the principles of faculty development was applied in the design of the instructional training materials used in this research.

Self-Instructional Design Guidelines

This section of the literature review includes information on self-instruction (S-I) as a design format. S-I was chosen because of its potentially strong complementary relationship with the instructional model chosen, the tenets of faculty development and adult learning and because S-I has been shown to be adaptable to dental education.

Self-instruction has many definitions and connotations; in this study, S-I will be employed as an educational method with the following characteristics: a) the subject matter is broken into small sequential units, b) the student is an active participant, c) immediate evaluation feedback is given, and d) self-pacing is provided for based on the student's ability (Gagne, 1977; Williams, 1981).

The characteristics of the self-instructional method stem from the postulates of the behavioristic learning theory (Williams, 1981); that is, the learner is rewarded for demonstrating a desired response to a particular task (stimulus). Classic programmed learning is a more formalized mode of self-instruction which clearly demonstrates the tenets of the behavioristic view. In S-I the student is presented with a statement containing a small amount of information, a question is asked based on that information, eliciting a response, and immediate feedback is given to the student, reinforcing the correct behavior. Complex learning occurs through a series of successful, reinforced trials.

Early advantages and disadvantages of S-I have been shown to be comparable to traditional methods of instruction. Students tend to have greater recall in addition to mastering materials in less time. Individualization can be promoted in different ways. Generally recognized disadvantages reported by Levine (1978) include the facts that developing materials often become mechanistic and that there may be a limit to the level of skills attainable through self-instruction. These negative characteristics were addressed during the development of the learning sequence of this study.

Controversy surrounds the effectiveness of using S-I to teach motor skills. Claims for (Mertens, 1971) and against (Norton, 1967) the applicability of S-I to motor skills can be found in the literature. No definitive research confirming either claim could be found.

After reviewing the general characteristics of S-I, the literature on self-instruction in dental education was surveyed. The use of S-I began in dental education in the 1960s and has continued to be a popular instructional mode through the 1980s. The reason for its popularity was accurately summarized by Williams (1981) in his review article, "Self Instruction in Dental Education from 1960-1980s." There was a "need perceived by dental educators and government agencies to dramatically increase the number of dental practitioners and to educate them as effectively and efficiently as possible" (p. 290). Self-instruction was selected because current research revealed S-I capable of increasing knowledge acquisition, bettering performance for lower ability students, increasing student motivation, reducing instructional time, and increasing knowledge retention.

William's investigation of the research since then revealed definite limitations. First, he found increased retention unsubstantiated. Second, he found specific characteristics of programs missing causing difficulty in discovering why success or failure occurred.

These two weaknesses of the dental research along with the general characteristics of S-I imply that self-instruction has been proven to be adaptable to dental subject areas. However, data are limited in the dental literature to support S-I effectiveness in teaching psychomotor skills. The subject of this research--calculus detection skills--is adaptable to all of the requisite features of the S-I method: the subject can be broken into small sequential units; the students (faculty) can participate actively, can receive immediate feedback, and can pace themselves. In addition, conservation of time is critical for busy faculty, and S-I promotes faster acquisition of knowledge.

Principles of Psychomotor Skill Learning

Conclusions from the literature on self-instruction lead to the next factor considered in establishing a faculty training program--characteristics of the task being taught. The overall aim of this study was to improve clinical evaluation; the method for achieving that goal was to refine calculus detection skills. Calculus detection is a behavior which requires precise manual activity and is termed a psychomotor skill (Bloom, 1956; Fleishman, 1962; Gronlund, 1978, Simpson, 1972).

No extensive research in experimental psychology pertaining to the acquisition and transfer of a motor skill is evident (Bilodeau, 1966, Hall et al., 1977; Irion, 1966, Mackworth, 1964; Parker and Fleishman, 1961; Smith, Wargo, Jones, and Smith, 1963; Wlanar, 1963). A summary of the above authors' conclusions as related to the acquisition and maintenance of calculus detection skills is

> 1. Immediate and continuous feedback produces the most rapid and efficient increase in proficiency; . . .

> 2. Signal detection (or calculus detection) can be facilitated by increasing the frequency of signals presented, providing feedback regarding performance accuracy and providing practice on the detection task. . . .
>
> 3. Feedback serves to improve the accuracy of detection performance and produces less decrement in performance as the time on task increases and facilitates performance when discriminations are difficult. .
>
> 4. Training with feedback transfers to situations in which feedback is not available
>
> 5. Skills that are practiced to a higher level of proficiency are retained for longer periods than those less well practiced. . . . (Hall et al., 1977, p. 367)

In addition to being a psychomotor skill, calculus detection is a task which requires tactile discrimination. The clinician (dental hygienist in this study) must rely on a sense of touch to direct movements of the dental instrument beneath the gingival tissue. The area for movement is very small (measured in millimeters) and not visible; also, a variety of stimuli cause a vibratory reaction of the dental instrument as it passes over the surface of the crown and root of the tooth. The clinician must be able to detect the surface irregularities and discriminate between anatomical structures and deposits which should be removed.

In a review of the literature on tactile discrimination learning by Troy (1971), the following factors were established. Discrimination ability is dependent on the individual's ability to perceive a change in the magnitude of the stimulus while simultaneously moving the instrument over the tooth surface. To encourage the development of an integrated sequence of activity, the learning situation for this skill (that couples discrimination with fine manipulatory movements) should resemble the actual clinical situation.

Mental practice improves both the speed and accuracy of response in a simple hand-eye coordinated task. Further, augmented feedback increases the ability to discriminate fine motor movements. Augmented is defined as precise feedback information, not just "correct/incorrect" feedback.

Conclusions from psychomotor and tactile discrimination literature provided additional guidelines for the development of the training program for faculty. The final subject area to be considered in the design of the training program was concept attainment.

Principles of Concept Attainment

Conceptual understanding and acquisition has been the subject of much research across a number of different content areas (Bruner, Goodnow, and Austin, 1956, Gagne, 1977; Joyce and Weil, 1980; Martorella, 1976; Piaget and Inhelder, 1969, Stahl, Sutton, and Corbett, 1975, Stahl, 1979, Spitzer, 1975). These individuals contend that the acquisition of concepts is critical for deep understanding of any content or activity. Recent research has led to the development of specific procedures for formulating lessons to facilitate concept learning (Merrill and Tennyson, 1977)

> A concept is assumed to be a set of specific objects, symbols, or events which share common characteristics (critical attributes) and can be referenced by a particular name or symbol. Concept learning is thus regarded as the identification of concept attributes which can be generalized to newly encountered examples and (used to) discriminate examples from non-examples. (Tennyson and Park, 1980, p. 56)

In several reviews of the research literature--Klausmeir and Feldman (1975), Stahl (1979) and Tennyson and Park (1980)--studies were cited which have led to the development of an effective instructional strategy for teaching concepts.

Many of the characteristics of the strategy were based on Gagne's (1977) original model of an eight-level learning hierarchy. A modification of Gagne's model was developed by Casteel, Buttes, Corbett, Gregory, LaCave, and Stahl (1974) to make it more applicable to social sciences, humanities and the arts rather than just to the pure sciences. The modified Gagnean model used only five levels of learning--verbal information learning, discrimination learning, concept-level learning, principle or rule learning and problem-solving learning.

Specific processes for the instructional strategy were developed over a period of time through the work of many researchers. Tennyson and Park (1980) reviewed over 60 studies which "directly led to the development of instructional design strategies for the teaching of 'classroom-like' concepts" (p. 65). Conclusions from their review were:

> According to the concept research literature reviewed, instructional stimuli should be selected and arranged to accentuate both the differences between the coordinate concepts and the scope of each concept. In addition, the instructional sequence should be organized in a contextual form, and the presentation order of the organized content should be decided according to updated information about the learner's knowledge state. A four-step process for concept teaching is proposed from the literature.
>
> 1. The taxonomical structure of the content should be determined. The three levels of concept structure superordinate, coordinate, and subordinate--should be analyzed with identification of critical and variable attributes.
>
> 2. A definition of the concept should be prepared in terms of the critical attributes, and a pool of examples should be prepared on the basis of critical and variable attributes.
>
> 3. The examples should be arranged in rational sets by appropriate manipulation of the attributes. Within a rationable set, containing one example from each coordinate concept, the examples should have similar variable attributes.
>
> 4. The presentation order of the rational set should be arranged according to the divergency and difficulty level among examples of the concept, and the presentation order of the examples within rational sets should be decided according to updated information about the learner's knowledge state. (p. 65)

This concept-learning strategy was selected as the instructional process for this research for several reasons. The learning requirement--calculus detection--is a highly complex skill. As has been previously discussed, calculus detection has a psychomotor component, a discrimination (tactile) component, and the investigator believes it also has a conceptual component requiring concept-learning tactics.

A concept has been defined as a set of specific objects, symbols or events which are categorized on the basis of common characteristics and which can be referenced by a particular label or symbol. Calculus detection was the label given to the event of interest. The abstraction of interest was smooth versus non-smooth and normal versus non-normal tooth surface. Conceptual understanding includes the ability to recognize diverse phenomena as being examples of the concept (Stahl, 1979). The diversities of tooth surfaces and deposits is as vast as the diversities of human beings (multiplied times 32 teeth).

The task of interest also goes beyond discrimination learning which requires a sorting process; but, of course, discrimination learning is a prerequisite skill. The faculty involved were assumed to have most if not all of the prerequisite skills including basic discriminatory skills for tooth surfaces.

Further explanation of the investigator's belief that calculus detection involved conceptual learning was based on the statement by Gagne that learning of a concept was not necessarily a verbal matter (1977). Verbal cues may be accompanied by other sensory cues. Calculus detection employs several kinds of stimuli as well as verbalization to assist the learner in developing the concepts

Summary

Clinical evaluation in dentistry is plagued with a history of minimal validity and low reliability (Abou-Rass, 1973; Baratti, 1974; Biller and Kerber, 1980; Fuller, 1972; Gaines et al. 1974; Hinkelman and Lang, 1973; Houpt and Kress, 1973; Hunter et al., 1975; MacKenzie, 1973; O'Connor and Lorey, 1978; Partridge and Mast, 1978). One of the evaluation tasks required is the discrimination between the presence or absence of unwanted deposits on a tooth surface which cannot be seen. The inference is made that if the detection skills of evaluators can be improved, the results will be improved evaluation. Some previous research on clinical evaluation in dentistry reported the effect of training on improving evaluation (Abou-Rass, 1973; Baratti, 1974; Fuller, 1972; Gaines et al., 1974; Goepferd and Kerber, 1980; Hinkleman and Lang, 1973, Houpt and Kress, 1973, O'Connor and Lorey, 1978); however, no studies were found which dealt with training faculty in an actual clinical skill.

After an unsuccessful search for an existing and effective training program, the investigator developed an aggregate approach. She employed as a guideline Gagne's model (1977), identifying the relationships between the phases of learning and the events of instruction. In addition, she developed an eclectic theoretical framework The subject areas included were faculty development, adult learners, self-instruction, psychomotor skill learning and concept attainment. She selected these components because of the unique requirements of improving faculties' calculus detection skills. She reviewed components of the model to identify the most current principles for application in the program design.

The literature revealed that faculty (adults) must see the training as needed and practical for their job, the training must be individualized, flexible and involve the learner; and it must not be threatening. Self-instruction was identified as the design format because it was amenable to all of the faculty needs and has been shown to be effective in dental subject areas.

After focusing on and identifying the needs of the target group, the learning task was considered. Calculus detection requires a psychomotor skill and the literature revealed that motor skill learning requires good feedback to facilitate transfer, and over-learning can increase retention levels (Hall et al., 1977). In addition, tactile discrimination was required and summary studies on tactile discrimination showed that mental practice as well as feedback were critical elements in training.

After the target population and the task were studied, instructional strategies were surveyed and concept attainment was selected as the method of choice. Research revealed that a specific series of steps had been developed which were readily adaptable to the task.

CHAPTER III

METHODOLOGY

Experimental Design

The objectives of this study were first to determine the level of validity in the calculus detection skills of faculty in dental hygiene programs and then to attempt to improve validity through a training program. It was assumed that the accuracy in calculus detection skills was low, and therefore, that the validity in clinical evaluation was low.

The basic research method of this investigation was a pretest-training-posttest process. There was one primary dependent variable--clinical evaluation skills of dental hygiene faculty. Validity in this case was defined as a measure of agreement with known criteria (i.e., the accuracy of calculus detection on a model of the human dentition using a dental instrument). The independent variable was a self-instructional training program.

Faculty participants were pretested twice on models to identify entry skill levels for calculus detection. Two methods of measurement were used to ascertain the entry levels--percentage of agreement and Cohen's Kappa measurement (Cohen, 1960). A self-instructional training module was given to the faculty; then, posttest data were collected immediately after training. McNemar's test for significant change was used (Seigel, 1956). The final phase of the research was a brief self-report survey to determine subjective data about the program.

Instrumentation

Primary instrumentation of this study consisted of a self-instructional training module containing brief written materials, a physical model of the human dentition, and self-grading forms (Appendix A). Other instrumentation included pretest and posttest forms and a self-report survey.

Development of the instructional package was of paramount importance in this study. Gagne's model (see p. 27, Figure 1) identifying the relationship between phases of learning and units of instruction was used (Gagne, 1977). Factors taken into consideration during the developmental process included principles of faculty development, characteristics of the adult learner, self-instructional design guidelines and principles of psychomotor skill learning and concept attainment (see p. 27, Figure 2).

The final phase of the research project was a self-report survey to gather information about the practice exercises, models and general opinions of the participants toward the project (Appendix B).

The Model

As previously stated, components of the models included a set of written introductory materials, a physical model of the human dentition and self-grading forms. Copies of each of the forms can be found in Appendix A. The models used are described below.

Each of the models had 28 teeth mounted in an arch relationship with simulated alveolar bone and gingival tissue. Teeth from the models were anatomical replications with roots and texture differentials simulating material of teeth as closely as possible. A representation of the alveolar bone was created with a dental stone material; gingival tissue was reproduced using a latex material, calculus was simulated with a mixture of materials which had varying granular characteristics. For additional information see the detailed instruction for making the models, Appendix C.

Utmost consideration was given to the creation of models with characteristics resembling the structures of the mouth to maximize transfer of learning from a simulation to actual patient care settings. Teeth were mounted in the arch with tilting, rotation and malalignment features. The alveolar arch created contained boney defects in combination with normal contours. Calculus was placed in amounts and locations commonly found in the mouth. Gingival tissues were contoured to look and feel like natural soft tissue with similar pliability and fragility, and the lubricating feature of saliva was reproduced through the use of a film of liquid soap applied to the teeth and gingiva prior to use

In addition to these features, the models were articulated and placed on bench mounts to simulate working positions. Cheeks were simulated by placing rubber dam material on the lateral side of the articulators. The tongue and hemorhagging were not feasible to simulate.

Selection of the dental instrument was the perogative of the participants. It was felt that their instrument of choice would facilitate transfer of learning.

A total of 25 models were fabricated. Six had identical characteristics and were used for the pretesting/posttesting, the other 19 had differing characteristics and were used by participants during the training sessions.

Instructional Module

Gagne's model (Figure 1, p. 27) was used as a guideline for developing the self-instructional module for improving calculus detection skills of faculty (Appendix A). The corresponding instructional activities were identified as a result of research and application of knowledge in five areas--faculty development, adult learning, self-instruction, psychomotor skill learning, and concept attainment.

As defined by the model, the motivation phase required an event which established expectancy in the learner. This was accomplished through the use of a pretest and a brief written summary on the importance of accuracy in clinical evaluation. Objectives of the training program were explained in writing prior to the pretest and were re-emphasized in the subsequent instructions to faculty. The instructions centered on relationships between clinical (calculus detection) skills and student learning. An effort was made to relate the learning task to efficacy in teaching.

Attention was directed to the learning event by the distribution of the module packages. These contained a set of dentition models, an explorer, soap (for saliva) and a packet of recording forms. Written instructions were also included to provide the participants with guidance on how to proceed through the module activities.

Actual practice of calculus detection abilities was the main activity of the module. Repetition was emphasized to enhance memory storage and retrieval and transfer of learning.

Participants were asked to continue to practice until a 90% (three surface areas) competency level was attained. Participants performed self-evaluation throughout all the practice activities and made the decision when to take a posttest; thus, they receive feedback continually throughout the instructional unit.

As stated previously, activities of the module were primarily instrumentation practice. Each participant was given a model and mount, a recording form, and a set of instructions (see Appendix A). Practice consisted of exploring designated teeth and recording findings. Then, by pulling up the gingival tissue and verifying presence or absence of deposits, each participant performed a self-evaluation. Each recording form initially had seven combinations of teeth with four surfaces each for checking (called practice subsets). This feature minimized memory interference and allowed the participants to practice repeatedly on the same model.

The second major feature of the instructional unit was the arrangement of the practice subsets. Each model and subset was developed following the principles of conceptual learning. Examples were arranged in rational sets according to the divergency and difficulty level of the detection task. Initial subsets had large, rough deposits located in easily accessible areas, whereas later subsets had small, smooth deposits in difficult-to-explore areas.

Additional features (related to faculty development and adult learning principles) considered during the design of the training program were

1. Administrative support was sought and received. A cooperative attitude was presented and resources were provided in terms of supplies and time.
2. The self-instructional format allowed faculty to direct their own learning in terms of time commitments.
3. Immediate and continuous feedback was supplied through the design of the model and the use of a self-check sheet. These features were incorporated to improve accuracy, reduce decrement in performance over time, and maximize transfer of learning.

Pilot Program

A pilot program was conducted in the fall of 1982 for several reasons. Of primary concern was the actual feasibility of using a set of models to measure and train individuals for calculus detection skills. The second concern of the pilot study was the assessment of the appropriateness and practicality of the data collection instruments. A third concern was the identification of any refinement of the hypotheses which might be required. Finally, the pilot study was used to obtain estimates of the time frame and the competency levels.

Participants for the pilot project were four Santa Fe Community College dental hygiene instructors chosen because they paralleled the research population. Professional backgrounds were varied and all had clinical teaching responsibilities.

Procedures

Each instructor was given two pretests using the models, a dental exploring instrument of choice and the scoring sheets. The pretests were administered individually by the investigator in a setting convenient to the participant. Feedback was provided on overall pretest scores, but no details were revealed about the errors made. This security served as motivation to achieve higher scores.

A meeting was set for discussing results, competency levels, time limits, and module activities. The participants were asked to read the module instruction, and request clarification when necessary. Notes were taken by the investigator for future revisions. The participants were asked to decide the level of desired competency to be achieved and to establish a time limit for completion. Finally, posttest appointments were made with the investigator. These considerations were implemented to provide participants with an opportunity for input.

After posttesting, data were analyzed using percentage of agreement. The data obtained were reported in Table 2.

Modifications of the research methodology as a result of the pilot study were:

1. The administration of the two pretests was accomplished at one time. Time requirements for exploring one model ranged from 12 to 25 minutes. Faculty felt it was too difficult to continue to repeatedly find blocks of the appropriate amount of time so two pretest dentition

Table 2

Pilot Program Results. Correct Responses and Individual Percentage of Agreement with Criteria

Faculty	Pretests 1 and 2 Correct Responses 1	Pretests 1 and 2 Correct Responses 2	Pretests 1 and 2 Percentage of agreement	Posttests 1 and 2 Correct Responses 2	Posttests 1 and 2 Correct Responses 2	Posttests 1 and 2 Percentage of agreement
1	81	89	75.9	101	92	86.2
2	75	82	70.1	103	104	92.4
3	80	85	73.7	96	107	90 6
4	87	91	78.5	110	108	97.3

models were developed to allow faculty to take both pretests concurrently.

2. The original scoring sheet for indicating calculus location was modified. Faculty felt the isolated subsets were too difficult to follow and led to many more errors due to exploring the wrong tooth. The modification of the scoring sheet was an adaptation of a full mouth dentition representation (see Appendix A).
3. Modification of the answer sheets led to the alteration of the practice exercises. Calculus deposits were placed, in rational sets (requiring easy then more difficult detection skills) in each quadrant so exploration could proceed in a more logical fashion as similar as possible to what would occur in a natural mouth.

Data collected from the pilot study indicated a lack of accuracy in the detection of calculus prior to training and an improvement following training. No statistical analysis was possible due to the small number of participants in the pilot. Pretest and posttest scores appeared to have a wide margin of difference (Table 2).

Percentage of agreement with the criterion was used to calculate individual validity. The criterion was correct identification of the presence or absence of deposits on the 112 possible surfaces of the teeth. Calculation of percentage of agreement (A) was accomplished using the following formula:

$$x = \frac{P_1 + P_2/2}{112}$$

where P_1 and P_2 were the number of correct responses from Pretest 1 and Pretest 2 respectively. Posttest scores were calculated in the same manner.

The final result of the pilot program was an estimate of the time needed to accomplish the pretests, training, and posttests. Participants were asked to record how much time was required to take the pretests and posttests. Results are reported in Table 3.

Table 3

Time Requirements for Pilot Program
Pretests and Posttests

Faculty	Pretests (1 & 2)	Posttests (1 & 2)
1	25 minutes	21 minutes
2	35 minutes	35 minutes
3	42 minutes	32 minutes
4	33 minutes	25 minutes

Summary

The pilot program indicated that the research project was feasible. Participants were able to improve their calculus detection abilities using simulated dentition models and a self-instructional training module.

Additionally, the pilot program resulted in identification of and several modifications in the pilot program. Pretests and posttests were given simultaneously to conserve time, recording forms were modified to achieve clarity, and calculus deposits were arranged to establish rational sets within quadrants. Time estimates were also established for pre- and posttesting.

Sample Population

The total population identified for this study was all faculty of dental training programs who were involved in the clinical evaluation of students' performances on calculus removal. The sample selected was 21 faculty of 3 dental hygiene programs--the University of Cincinnati, Pensacola Community College, and Tallahassee Community College.

Selection procedures were based on availability and willingness to participate. The investigator recognized the critical statistical limitation due to the lack of a randomized sampling procedure and of a control group. This limit was considered an asset to the motivational aspect and success or faculty development programs.

Other factors influencing the sample were the time and financial constraints of the investigator.

Analysis of Data

The first question of this study was, to what extent can dental hygiene faculty judge whether calculus is present. The investigator believed that accuracy was low (i.e., that faculty judgments were not consistent with criteria). The first hypothesis of the study developed from this question and was stated in the null form:

> H_o1: Dental hygiene instructors as individuals are able to accurately identify calculus when criteria are available for comparison.

The method of the measurement used to determine whether to reject the null hypothesis was percentage of agreement. Results from this statistic identified the faculties' agreement with the criteria (i.e., did faculty accurately identify the presence or absence of calculus)? Acceptable values for percentage of agreement were defined by the author as being greater than 90%.

> H_o2: Dental hygiene instructors as a group are able to accurately identify calculus when criteria are available for comparison.

Two methods of measurement were used to test this hypothesis. Percentage of agreement measurement was used again, determining group response rate using the means of percentage of agreement for individuals. The second measurement used was Cohen's Kappa (Cohen, 1960), a chance-corrected statistic for assessing interrater agreement using a nominal scale classification.

As stated in the preceding paragraphs, percentage agreement was used in two ways (1) to find agreement of the individuals' with criteria and (2) to find agreement of the group as a whole when compared with a criterion. The criterion was the actual presence of calculus on the surface of a tooth. Calculus was placed on the surfaces in a random fashion to assure sample independence.

Calculations for percentage of agreement were done in the following manner. Each subject was given two pretests P_1 and P_2 and the correct responses were recorded for each session. A total of 28 teeth with four surfaces might have had calculus resulting in a possible total of 112 correct responses. Percentage of agreement for the individual was computed using the following formula:

Percentage of Agreement for the Individual (A_1)

$$A = A_1 \frac{(P_1 + P_2)/2}{112} \times 100$$

Percentage of agreement for the group, all faculty participants, was computed using the mean percentage of agreement scores.

Percentage of Agreement for the Group (A_g)

$$A - A_g \quad \frac{(A_1 + A_2 + t. \quad Ax}{X}$$

Acceptable values for percentage of agreement were defined by the investigator as being greater than 90%. At least 50% agreement could occur by chance, and the task had importance to warrant a high expected level of agreement.

In addition to computing percentage of agreement, Cohen's Kappa for measuring interrater agreement was selected to further control for chance when using nominal data. The formula used was $K = P_o - P_c$ where P_o represents the observed proportion of interrater agreement attributable to chance factors (P_c) alone.

Cohen's Kappa was selected because recent literature has argued against using percentage of agreement when using nominal scales (Hartman, 1977), because percentage of agreement has provided misleading information using nominal scales. Nominal scales define specific categories by name and all observations can be categorized. The measurement of interest in hypothesis 2 is criterion-referenced and classifies examinees into mastery and non-mastery categories.

The conditions of the research met the following assumptions for using Cohen's Kappa:

1. The categories used must be mutually exclusive and nominal. The presence or absence of calculus was certain and therefore mutually exclusive and exhaustive of possible alternatives.
2. The cases classified must be independent. Calculus was placed on the 112 tooth surfaces randomly, thus the cases classified were independent.
3. Raters must operate independently in the classification; each rater was detecting calculus without interaction with other participants.

Results of Cohen's Kappa statistic provided a coefficient estimating the amount of shared variance in the group for calculus detection skills. The amount of error due to chance was 50% because there were two choices, calculus or no calculus. If there was no agreement except for that due to chance, then the coefficient would equal zero, if there was total agreement with the criteria all the time the coefficient would be one. The acceptable value for the coefficient was set by the investigator as .90 or greater. This coefficient reflects the level of agreement the group has with the criteria.

After establishing the low level of accuracy of calculus detection skills, the representatives at each school administered a training program designed to improve the skills of the participants. A posttest was administered to the faculty after completion of the training program and a comparison was made with the pretest scores. The third hypothesis reflects the interest in this comparison:

> H_o3: Training will not significantly improve the ability of dental hygiene instructor to detect calculus

McNemar's test for significance of changes, a statistical nonparametric test for nominal measurements, was used to test the null hypothesis. McNemar's test may be applied to a pretest/posttest design in which each subject is self-controlled and in which two related samples are used

"To test the significance of any observed change by this method, one sets up a fourfold table of frequencies to represent the first and second sets of responses from the same individuals" (Seigel, 1956, p. 63). The table yields measures of expected frequency of change in cases measured. This value is used in the following formula to compute the significance of change.

$$^2 = \sum_{A,D} \frac{(O - E)^2}{E}$$

Procedures

Packages of models and instructions were mailed to each of the three dental hygiene programs in January, 1983. The project directors at each school were given instructions on administration of the testing and training (Appendix A). Each director was asked to return the pretesting, posttesting and survey results by May 15, 1983.

CHAPTER IV

PRESENTATION OF FINDINGS

Introduction

Clinical evaluation of dental procedures has been shown to be inaccurate and inconsistent. The literature revealed a limited amount of research and inconclusive research results. This study investigated one aspect of clinical evaluation in an attempt to improve the accuracy of evaluation. The skill chosen for evaluation was calculus detection, the population was dental hygiene instructors, and the variable of interest was training.

Instructors from three different dental hygiene schools participated in the study. The research design was a pretest-treatment-posttest using simulated dentition models and a self-instructional training module. The dentition models were designed to closely approximate the conditions found in a mouth which has moderate to advanced periodontal disease (20-50% bone loss) and which was just previously been cleaned by a student. The calculus deposits were small and difficult to find to represent the same evaluation task instructors must perform clinically. The self-instructional training module was specifically designed according to the current research on adult learning, faculty development, psychomotor skill learning and concept attainment. Each participant was pretested and posttested using parallel dentition models and individualized training activities.

Outcomes were obtained through statistical analysis of the pretests and posttests and through self-report survey. The statistics used for analysis of the data (Percentage of Agreement, Cohen's Kappa and McNemar's Test for Significance of Change) were chosen to control for variables related to criterion-referenced measurements. The survey was used to obtain information about the training exercises and about the participants' subjective reactions to the project.

This chapter was organized around the three hypotheses stated in the null form and the surveys. The data are presented following each hypothesis. Conclusions are found in Chapter V.

Hypotheses 1

H_o1: Dental hygiene instructors as individuals are able to accurately identify calculus when criteria are available for comparison.

Twenty-one dental hygiene faculty from three dental hygiene schools completed two pretests which measured the ability to detect calculus. The calculus was placed on a random sample of teeth and a key was made to identify its location. The location of the calculus became the criterion by which the faculty were evaluated. A correct response meant that the faculty either found and recorded a calculus deposit when it was present or did not record calculus when it was not present.

The accuracy value for percentage of agreement was defined by the investigator as being greater than 90%. This value was chosen for two primary reasons. At least 50% agreement could occur by chance because there were only two choices, either presence or absence of calculus. Second, the task of calculus detection is sufficiently important to warrant a high expected level of agreement since students are judged competent or incompetent based on this criterion and a patient's state of periodontal health depends on detection and removal of calculus.

Results of each individual participant's pretests are reported in Table 4. Correct response scores for both pretests (P_1 and P_2) are reported; the percentage of agreement measurement is an averaged score of the two pretests.

Only one of 21 participants achieved an acceptable accuracy, greater than 90, score. This provided evidence that the dental hygiene instructors tested were inaccurate in detecting calculus deposits.

The significance of these data relates to the issues of validity (and reliability). In clinical settings where instructors are required to evaluate student performance, the presence or absence of calculus deposits can rarely be confirmed. Instructors may develop a degree of consistency

Table 4

Pretest Scores of Individual Faculty for Calculus Detection Reported as Percentage of Agreement

$$\left(A = \frac{P_1 + P_2/2}{112}\right)$$

Faculty	Correct Response P_1	P_2	Percentage of Agreement
A	78	53	58.5
B	91	95	83.0
C	87	92	80.0
D	94	89	81.7
E	71	96	74.6
F	89	96	82.6
G	94	100	86.6
H	102	93	87.0
I	97	105	90.2
J	82	99	80.8
K	96	92	83.9
L	80	91	76.3
M	96	95	85.3
N	84	94	79.5
O	91	94	82.6
P	90	100	84.4
Q	91	97	83.9
R	81	93	77.7
S	71	88	71.0
T	74	97	76.3
U	95	104	88.8

n = 21

individually or with other instructors, but this reliability may not be valid. One instructor may repeatedly miss calculus on a maxillary molar incipient furcation remaining consistent but not correct. Several instructors may explore a patient's teeth and agree on the results, but all may have missed a deposit on the mesial concavity of a maxillary first premolar. An opportunity to obtain reliable and valid feedback on an individual basis is provided by this research.

The statistics indicate that only one in 21 instructors tested could accurately determine the presence or absence of calculus when accuracy was defined as a consistent 90% correct response. Dental hygiene instructors were unable to accurately identify calculus when criteria were available for comparison; therefore, H_o1 was rejected.

Hypothesis 2

H_o2: Dental hygiene instructors as a group are able to accurately identify calculus when criteria are available for comparison.

Two methods of measurement were used to test this hypothesis. Percentage of agreement measurement was used to determine whether the faculty as a group was able to accurately identify calculus and group response rate was determined by using the means of percentage of agreement for individuals (see Table 4). The second measurement was

Cohen's Kappa, a chance-corrected statistic for assessing interrater agreement using a nominal scale classification.

Percentage of agreement was used to determine the agreement of the group as a whole when compared with criteria. As with the first hypothesis, the criterion was whether calculus was actually present or absent on the surface of a tooth. The calculus was randomly placed on the surfaces to ensure sample independence.

Computation of percentage of agreement for the group (A_g), all faculty participants, was accomplished using the following formula,

$$A_g = \frac{A_1 + A_2 + \ldots A_x}{x}$$

where A_1, A_2, . . . represent the individual percentage of agreement scores and x equals the number of participants.

The computed percentage of agreement for the group was 80.7 and accuracy was defined as being greater than 90%. Dental hygiene instructors as a group were unable to accurately identify calculus when criteria were available for comparison, therefore H_o2 was rejected.

In addition to percentage of agreement, Cohen's Kappa was used to test the second hypothesis. This statistic was chosen because control for chance when using nominal data is achieved and the assumptions for using Cohen's Kappa are met. The categories, presence or absence of calculus, are mutually exclusive and nominal and the cases classified were randomly selected and were, therefore, independent. The participants operated independently by detecting calculus without interacting with other participants.

Participants in this research were asked to identify the presence or absence of calculus on 112 tooth surfaces. Only two choices, present or absent, were available. The actual presence or absence was known and therefore valid. An instructor or group of instructors were considered to be masters or non-masters determined by the ability to agree with the criteria. Mastery level was defined as 90% agreement with the criteria as explained previously.

The measure produced from the statistic was a measure of agreement between the decisions in repeated administration. The instructors were the measuring instrument and instructor validity was measured. The coefficient produced provided an estimate of the participants' abilities to repeatedly respond correctly. If the coefficient was low, instructors could not repeatedly find calculus. If the coefficient was high, instructors could consistently find calculus.

$$K = (p_o - p_c)/(1 - p_o)$$

where p_o was the observed proportion of agreement, p_c was the expected proportion of agreement, and K was the proportion of agreement that existed over and above that expected by chance alone.

Table 5 provides summary information for the sample data on the classification of examinees into two mastery states on two test administrations. The numbers in the matrix are proportions of individuals which fall into the designated master and non-master categories on Pretest 1 and 2. Computations for obtaining Kappa are below the matrix.

The upper limit of K is +1 and may only occur when the marginal proportions for different administrations are equal. The lower limit of K is close to -1 indicating extreme decision-making inconsistency (i.e., the observed agreement is less than expected by chance). Significance testing for K is of little value (Cohen, 1960); "to know merely that K is beyond chance is trivial since one usually expects much more than this in the way of reliability. . ." (p. 44)

The computed coefficient was -.09; therefore, H_o2 was rejected. The conclusion was as hypothesized; dental hygiene instructors as a group were unable to accurately detect calculus when a criterion was available for comparison.

Hypothesis 3

H_o3: Training will not significantly improve the ability of dental hygiene faculty to detect calculus.

Table 5

Classification Matrix for Cohen's Kappa

Pretest 2 Pretest 1		Mastery States		Marginal Proportions
		Master	Non-master	
Mastery States	Master	0	.05	.05
	Non-Master	.19	76	.95
Marginal Proportions		.19	.81	

$P_o = 0.0 + .76$

$= .76$

$P_c = (.05)(.19) + (.95)(.81)$

$= .78$

$k = (.76 - .78)/(1 - .78)$

$= -0.09$

The third hypothesis was tested after participants were asked to complete a self-instructional training program. Each faculty was posttested on the same two dentition models used for pretesting. The results are presented in Table 6. Correct response raw scores for posttesting were computed, then averaged, to obtain individual percentage of agreement scores. Individual and group percentage of agreement scores for posttesting were computed in the same method as for pretesting.

Table 7 is a comparison of each participant's pretest and posttest percentage of agreement scores. The group scores are summarized following the individual scores. For example, instructor A agreed with the criterion 58.5% at pretesting and improved to 68.5% agreement at posttesting. The instructor did not obtain mastery as defined by the investigator.

McNemar's test for significance of changes was used to test the third hypothesis. A nonparametric test, McNemar's test for significance of changes, was chosen because it was designed to determine if a significant change has taken place over time. In this research the statistic was used to determine if the instructors as a group had improved significantly in calculus detection skills. The mastery level of greater than 90% agreement with the criterion was again used to categorize instructors into mastery and non-mastery divisions.

Table 6

Posttest Scores of Individual Instructors for Calculus Detection Reported as Percentage of Agreement

Instructor	Correct Responses P_1	P_2	Percentage of Agreement
A	69	95	68.8
B	91	98	79.9
C	97	88	82.9
D	104	105	93.3
E	99	109	90.2
F	101	105	92.0
G	99	105	91.1
H	104	104	92.9
I	109	102	92.9
J	99	99	83.9
K	92	105	87.9
L	92	93	82.9
M	94	100	96.9
N	89	105	86.9
O	102	104	92.0
P	99	103	90.2
Q	105	82	83.5
R	102	103	91.5
S	90	104	99.9
T	104	99	90.2
U	109	103	93.3

Table 7

Comparison of Individual Pretest and Posttest Scores

Instructor	Percentage of Agreement Pretest	 Posttest
A	58.5	68.8
B	83.0	79.9
C	80.0	82.6
D	81.7	93.3*
E	74.6	90.2*
F	82.6	92.0*
G	86.6	91.1*
H	87.0	92.9*
I	90.2*	92.9*
J	80.8	83.9
K	83.9	87.9
L	76.3	82.6
M	85.3	86.6
N	79.5	86.6
O	82.6	92.0*
P	84.8	90.2*
Q	83.9	83.5
R	77.7	91.5*
S	71.0	86.6
T	76.3	90.2*
U	88.8	93.3*
Group Totals	80.7	87.6

*The asteriks indicate the participants who achieved mastery level.

Computations for McNemar's Test were accomplished by using a two-way matrix in which individual performance at each test period were categorized. Table 8 illustrates the four categories for classification; for example, A represents the participants who were masters on the pretest but were non-masters on the posttest.

Table 8

McNemar Data Summary Matrix

Pretest	Posttest Non-master	Master
Mastery	A	B
Non-mastery	C	D

This table yielded measures of frequencies representing the pretest and posttest results from the same group of individuals. In using this statistic the interest was in cases which showed change. The sums of the cases in cells A and D represented the total number of people who changed (i.e., went from master to non-master or non-master to master). The formula used to compute the test value was

$$\chi^2 = \sum_{\text{cells}\,-\,A,D} \frac{(O-E)^2}{E},$$

where, O = observed number of cases in the i^{th} category and E = expected number of cases under H_o in the i^{th} category. Since in the McNemar test for significance of changes the interest was only with cells A and D, the formula can be expanded and collected as follows:

$$\chi^2 = \frac{(A-D)^2}{A+D}, \text{ with df} = 1$$

Substituting the measured values into the matrix in Table 9 yielded the measures of expected frequencies of change from the categories master and non-masters.

Table 9

Summary of Data for McNemar Statistic

	Posttest	
Pretest	Non-master	Master
Master	0	1
Nonmaster	9	11

Substituting into the above formula yielded a computed value of 11 with p = .001 (less than .05) and df = 1; leading to the rejection of H_o3. The conclusion was that instructors showed a significant tendency to change from nonmastery to mastery level in calculus detection skills from the pretest to the posttest

Survey Findings

A follow-up questionnaire was sent to all instructor participants after completion of the posttests (Appendix B). The objective of the survey was to obtain additional information about the training program and the dentition models. The training program was a self-instructional module based on Gagne's theoretical model (Figure 1, p. 28) and the knowledge bases of five subject areas--faculty development, adult learning, self-instruction, psychomotor skill learning and concept attainment (Figure 2, p. 29). The dentition models were specifically designed and created for this research to replicate the structures, textures and contours found in the human oral cavity.

The first section of the questionnaire was designed to obtain information about the training exercises completed by the participants. The module directions indicated that each instructor should practice exploring a series of different dentition models and, after completing a quadrant, should self-evaluate and continue with the module. Instructors were to determine individually the amount of practice needed based on the ability to consistently find calculus. Participants were encouraged to practice at convenient times in accordance with work schedules.

Survey results indicated that 11 of the 21 participants completed a series of practice exercises. The amount of time spent on the practice exercises ranged from 45 minutes to two hours, five used soap as instructed, seven followed the order indicated, and five self-evaluated calculus detection after practicing. Seven practiced using four models, two practiced using three models, and two practiced using two models.

The remaining 10 participants reported doing no practice exercises. The most common reason for not doing practice exercises, as reported by nine participants, was lack of time in the teaching schedule.

The second part of the survey was designed to collect more information on the models. No pattern in the recommendations for improving the models emerged. The responses were obtained in two ways--a checklist and an open ended request for suggestions. The check list appeared as illustrated below with the numbers in parentheses equalling the number of participants checking the statement.

I feel like the project would have improved my evaluation skills if there was (were)

(2) more bone loss on the models

(1) less bone loss on the models

(7) more calculus on the models

(1) less calculus on the models

(7) no gingiva for some of the practice time

Four respondents suggested a need for improved calculus retention. Four suggested improvement of the epithelial attachment apparatus. Four suggested a need for improvement in the gingival texture.

Participants were also asked to make suggestions about project improvement. Three said an introductory seminar was desirable, and nine wanted specific supervised practice times.

Additional Findings

The scoring method for pretesting and posttesting yielded additional data of interest--time requirement for examination, types of errors and differences in the professional backgrounds of instructors. Summaries of this information are presented in the following paragraphs.

Time Requirements

Participants were asked to record beginning and ending times for the pretests and posttests. Table 10 contains those times. The pretest mean time was 35 minutes, the posttest 24 minutes. The average time required for a thorough examination of two full dentitions by instructor was 30 minutes. This information holds potential value for those planning schedules requiring calculus detection skills such as teaching institutions, licensing examinations, and private practice scheduling of patients for re-examination.

Table 10

Averaged Time Requirements for Pretests 1 and 2 and Posttests 1 and 2 for Individuals and as a Group

	Individual	
	Time in Minutes	
Instructor	Pretests	Posttests
A	34	23
B	22	20
C	30	20
D	35	10
E	34	13
F	45	30
G	35	--*
H	30	--*
I	45	45
J	24	20
K	45	10
L	35	35
M	45	35
N	46	20
O	36	25
P	33	21
Q	41	25
R	33	21
S	28	25
T	42	35
U	22	23

	Group		
	Pretest	Posttest	Overall
Mean	35	24** minutes	30

* times not reported

**n = 19

Types of Errors

The hypothesis testing of the research was based on correct response of the participants, but error scores were also tabulated. Table 11 presents a summary of those results. Two types of errors were considered, false negatives (FN) and false positives (FP). A false negative indicates presence of unfound calculus deposit and a false positive indicates absence of reported calculus deposit

Recognition of the types of errors made by the faculty was important for several reasons. A false negative was judged by the investigator to be the more critical error since, in the clinical setting, undetected calculus deposits will subsequently lead to a continuation of the disease process. A false positive is not necessarily harmful to a patient but has implications in the evaluation process. A student's competency could be misjudged based on false positive responses. As seen in Table 11, a greater number of false positive errors occurred and the range of scores was large--26, 33, 21 and 22.

In addition to these group measurements, interesting error changes occurred and trends for error changes could be noted within the group. Individual feedback would be beneficial to the participants in this study. This information has potential use in improving validity in calculus detection skills.

Table 11

Summary of Errors

Faculty	Pretests				Posttest			
	1		2		1		2	
	FN	FP	FN	FP	FN	FP	FN	FP
A	11	23	26	36	6	17	4	23
B	17	4	9	6	8	13	5	19
C	13	12	6	14	9	6	4	20
D	18	0	5	18	8	0	6	1
E	12	29	10	6	13	3	6	0
F	12	11	3	13	8	3	4	3
G	13	5	6	6	9	4	5	2
H	8	2	6	13	2	6	4	4
I	11	4	4	3	5	11	2	8
J	17	13	1	13	15	8	5	8
K	11	5	5	15	18	2	5	1
L	10	22	4	17	17	3	6	13
M	11	5	7	10	13	5	10	2
N	7	21	4	14	15	8	2	5
O	1	20	1	17	4	6	3	5
P	6	16	6	6	7	6	6	3
Q	7	14	4	11	9	21	2	5
R	14	17	8	11	3	7	5	4
S	15	26	7	17	11	10	3	5
T	12	26	7	8	4	4	7	7
V	11	6	5	3	6	0	6	3
range	17	26	25	33	16	21	8	22
mean	11	13	6	12	9	7	5	7

Professional Backgrounds of Faculty

Although pretest and posttest scores were anonymously submitted to the investigator, information was collected on the professional backgrounds of each participant. Possible categories were 1) dental hygienist--full-time instructor, 2) dental hygienist--part-time instructor, 3) dentist--full-time instructor, and 4) dentist--part-time instructor. The distribution was as follows:

1. dental hygienist--Full time	9
2. dental hygienist--Part time	8
3. dentist--Full time	1
4. dentist--Part time	3

Only one notable relationship to the professional backgrounds of the participants and scores emerged: part-time dental hygienists scored higher. Inadequate data were available to rely on these results and this relationship is noted only as a trend.

This information has potential use in relation to faculty development. Additional statistical information could have an influence on the type and extent of future programs to achieve validity in clinical evaluation. In addition, general error trends should be studied to aid in instructional designs.

Summary

Analysis of the data yielded the following results:

1. Twenty of 21 dental hygiene instructors tested were unable to achieve a 90% accuracy or percentage of agreement score when asked to identify calculus.
2. Dental hygiene instructors as a group were unable to accurately identify calculus at a 90% level of accuracy. The group percentage of agreement score was 80.7. Cohen's Kappa statistic indicated extreme inconsistency and a negative value was obtained, - 0.09.
3. The validity of dental hygiene instructor calculus detection skills improved significantly from pretest to posttest.
4. Self-report surveys indicated a wide variety in the amount of time spent in practicing calculus detection.
5. The average time to complete the pretesting was 35 minutes; the average time for posttesting was 24 minutes.
6. More false positive errors than false negative errors occurred on both pretesting and posttesting.
7. Only one relationship appeared between instructor calculus detection skills and professional background; part-time dental hygiene instructors demonstrated a tendency to obtain higher scores.

CHAPTER V

DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

Introduction

Clinical evaluation of student performance has been recognized as an essential and complex task in dental hygiene education. The standards used to judge clinical competence have traditionally been highly subjective with invalid results. The consequences of invalid clinical evaluation affects patients, students, and instructors. Patients may receive substandard care, students become frustrated, and instructors function inefficiently.

Research and literature on the issues and problems associated with clinical evaluation is scarce. However, some information is available on variables requiring further research. This study was designed to investigate one of the variables affecting clinical evaluation--training of instructors. This study sought to improve instructor calculus detection skill through the use of a training program.

Instructors in three dental hygiene programs participated. The skill of interest was calculus detection, and the training program was a self-instructional module. Dentition models were developed specifically for use in the research. Pretests were administered to establish individual skill levels and group skill levels. The training program was developed using Gagne's instructional theory and current research from five subject areas (adult learning, self-instruction, psychomotor skill learning, concept attainment, and faculty development). Identical posttests were administered following training and comparisons were made to determine if the instructor calculus detection skills improved significantly.

Postulating low accuracy in calculus detection skills and improvement in these skills through training resulted in three hypotheses to be tested. The pretesting and posttesting were criterion-referenced and the methods of analysis chosen were appropriate to criterion-referenced testing. Calculus deposits were placed on the teeth of dentition models, location of the calculus was randomly selected, and the testing models were standardized for each school. The criterion applied was the actual finding of the calculus. No question of criterion validity arose because the calculus was visible if the simulated gum tissue was removed. Acceptable accuracy was determined by the investigator to be 90% agreement with the criterion.

Using these methods of testing and analysis, the results indicated that the dental hygiene participants were initially inaccurate at the 90% or above level in calculus detection abilities; however, after some training, instructors improved in accuracy significantly at posttesting.

The remainder of this chapter contains the major findings and conclusions of the research findings and recommendations for further study.

Discussion and Conclusions

The dearth of research and consistent conclusions about clinical evaluation in dental education prompted this study. Personal experience of the investigator, informal discussions with other dental hygiene instructors, and periodontal research reports all indicated that instructor calculus detection skills were low. Calculus detection is a primary skill requiring dental hygienist and dentist competency development. Teaching and evaluating competency of calculus detection skills requires accuracy in those skills to provide meaningful feedback and make valid judgments about student progress.

Hypothesis 1 and 2

H_o1: Dental hygiene instructors as individuals are able to accurately identify calculus when criteria are available for comparison.

H_o2: Dental hygiene instructors as a group are able to accurately identify calculus when criteria are available for comparison

Verification of the lack of validity in instructor calculus detection skills generated the first two hypotheses of this study. The outcomes of the pretesting substantiated the lack of accuracy in instructor calculus detection skills on both an individual and a group basis. Only one person of 21 achieved a greater than 90% accuracy. The percentage of agreement for the group was 80.7 and accuracy was defined as being greater than 90%. In addition, Cohen's Kappa was computed to establish a chance-corrected coefficient; K was -0.09 indicating extreme inconsistency in the calculus detection skills tested.

Efforts to eliminate influencing variables included development of the dentition models with features to control for accuracy and standardization, knowing the location of the deposits, and keeping the anatomical features of the mouth consistant. Conditions for pretesting are specifically outlined in Appendix A, and two pretests were administered.

In addition, the statistical methods employed for analysis were selected based on the dictates of the testing situations. Validity was tested using a criterion-referenced measure which controlled for a chance factor and a rare event. Examinees could achieve 50% agreement with this criterion by guessing. The amount of calculus present was small compared to the number of tooth surfaces to be explored and an examinee could achieve a high percentage of agreement score without identifying any calculus deposits. Cohen's Kappa was employed to reduce these influencing factors.

Hypotheses 1 and 2 were generated and reported separately because the results of the pretesting were intended as motivating factors for the instructor participants. Individual results were to be given to each participant and group results were to be provided as further confirmation of the need for improvement.

The results of this pretesting clearly indicated a lack of accuracy in the instructors' calculus detection skills.

Limitations to the results must be recognized. Although the literature supports the use of models in educational settings, the transferability of the skills to be learned is not guaranteed. While every attempt was made to develop models that replicated the features of a human mouth, differences existed which may have influenced the results. Repeated usage of the model possibly altered its features creating another limitation. Participants reported calculus and plaster had flaked off some of the models. A third limitation of the models used for this study occurred due to unfamiliarity of the instructors with this new product. Because the faculty had never explored this particular type of model the accuracy scores may have been unduly low. The use of two pretest scores was employed to partially control for this limitation.

The investigator recognized the small sample size and lack of random sampling procedure. These features were unavoidable but the influence on the results must be considered.

Hypothesis 3

H_o3: Training will not significantly improve the ability of dental hygiene instructors to detect calculus.

After assessing the calculus detection skills of the participating dental hygiene instructors, a training program was provided to improve the skills of the instructors. Posttests were administered and change was measured. The results indicated that instructors significantly improved in calculus detection skills. At the .05 confidence level the null hypothesis was rejected

As with the pretesting, efforts were made to control for existing variables. Two posttests were administered using the same dentition models. Statistical analysis was accomplished using McNemar's test for significance of change because of its applicability to "before and after" designs and nominal data.

Some of the same issues identified in pretesting as questionable existed with the posttesting. Transferability, alteration of the models through usage, and sampling procedures restricted the conclusiveness of the results of this phase of the research. Several other limitations must may be recognized. Improvement in calculus detection skills could have been a function of repeated evaluations of the same models. The investigator believed this limitation remote because of the number of tooth surfaces involved, the minimal number of deposits, and the length of time between pretesting and posttesting. At least three to four weeks passed between the pre- and posttesting. The final limitation recognized was related to the variability in training. In those cases where no training was reported, an improvement in calculus detection was reported in some instances. The question arises as to improvement solely related to familiarity with the models and the training effect of pretesting and posttesting.

Survey Findings

Follow-up data were sought from the dental hygiene faculty participants in order to gain a more comprehensive understanding about the training program and the dentition models. The results of the survey indicated that the training which took place was diverse and the models could be improved. The training program was designed to be a self-instructional module. This format was selected because of its complementary relationship with the theories of adult learning and faculty development. The activities of the module were designed to accommodate the postulates of psychomotor learning and concept attainment, following the guidelines of Gagne's instructional sequence. Module instructions were provided to the participants in a brief and concise format and reinforcement of these instructions was provided to the project administrators through discussions and written directions.

Reasons reported for diversity in the training program included a) not enough time in the teaching schedule, b) lack of motivation, and c) inconvenience.

These findings led to the following conclusions:

1. The training module employed in this research may have been only one of many variables influencing performances.
2. Time constraints create major barriers to faculty development.
3. Stronger motivators are required to prompt and maintain desired practice activity levels.

Diversity in the survey results regarding the models also existed. Although the participants did not identify any one particular outstanding negative characteristic, the investigator believes the mode of calculus attachment needs to be improved. Participants noted two additional modifications that might improve response 1) establish a more defined epithelial attachment and 2) adjust gingival texture.

Additional Findings

In addition to the hypotheses testing and survey results, this study yielded data on three other topics of interest to calculus detection and dental hygiene clinical evaluation, a) time requirement for completion of an evaluation for calculus, b) types of errors made when exploring for calculus, and c) relationships between professional background and calculus detection skills.

The average time for two full-mouth explorations for calculus on the models was 30 minutes. Although not conclusive, a strong indication exists that at least 15 minutes should be allowed for the exploration of a full dentition to determine the presence or absence of calculus.

This information has implications for scheduling of faculty and students in educational settings, time allowances in licensing examinations, and scheduling of patients in private practice settings. The amount of time available to students and faculty in educational settings is crucial to learning. If sufficient time is not available to students and faculty for evaluation and feedback, efficiency and effectiveness are reduced. Licensing examinations in dentistry require calculus detection and removal. Decisions for licensing a professional are based on an examiner's ability to detect calculus. If the time for evaluation is not sufficient, the decision to license or not license will be deleteriously altered. In private practice settings calculus detection is necessary. The patient's periodontal health is at risk. Sufficient time must be allotted to evaluate the patient in order to ensure calculus removal. The time required to detect calculus for the reasons listed above was noted and recommendations will be made regarding these findings.

A second topic of interest, not included in the research but warranting attention, is the type of errors made by the instructors participating in the study. Two types of errors were possible--a false positive and a false negative. The false negative error was the more critical of the two errors. If calculus deposits were present and not detected, disease would continue. A false positive error, absence of reported calculus deposits, has implications for evaluation but is not detrimental to the health of a patient. Knowledge of the types of errors would also be useful in an instructional setting. Feedback to faculty and students about errors made influences the learning activities selected to improve those skills.

The summary of the types of errors and error frequency revealed the following: 1) more false negative than false positive errors, although the range in errors was wide, 2) more research is needed to identify and reduce the kinds of errors made during calculus detection, and 3) location errors should be analyzed as to tooth, surface, and depth.

The third topic of interest regards the relationship between professional backgrounds of dental hygiene instructors and calculus detection skills. Opinions in the dental profession regarding personnel possessing acceptable levels of skills to perform certain tasks. The backgrounds represented in this sample distribution were divided into four categories. The results indicated a tendency for the part-time dental hygienists to score higher. Based on this data alone, definitive conclusions could not be made, only tendencies identified. The need for more research is apparent.

Summary of Conclusions

The validity of dental hygiene instructor calculus detection skills tested in this research was minimal. These skills improved significantly from pretesting to posttesting. Several factors and uncontrolled variables may have influenced the research results. Determining whether the demonstrated increase in skill levels reflected real differences or limitations of the study design and whether these differences were due to the training or some other variable was difficult.

Recommendations for Further Study

The findings of this study suggest that training has an influence on calculus detection skills and subsequent clinical evaluation in dental hygiene education. The results were not conclusive and the subject requires more extensive research. This section contains recommendations of areas for further study, including specific suggestions related to the dentition models, training program, and statistical analysis.

Dentition Models

1. Greater effort should be made to develop models simulating the conditions found in the oral cavity. Gingival tissues, osseous contours, and calculus deposits should be improved in future models.
2. Transference of skills from models to clinical conditions should be verified. Studies should be designed to determine if calculus detection skill levels are the same on the models and in the mouth, and if improvement of skills on the models indicates improvement of skills in the mouth.

Training

3. Control of training should be developed and alternative methods of training should be tested. Variables in the self-instructional package and other training formats should be tested with control groups to determine the most effective methods.
4. Motivational factors for instructors should be identified and investigated. Several possibilities include academic credits, continuing education hours, and release time.

5 The effects of time variations in training should be scrutinized. Does the amount of time spent in training affect the skill levels obtained?

Statistical Analyses

6. The study should be replicated with an effort to increase sample size to determine the reliability of results and of the evaluators.
7. Additional efforts should be made to find and develop adequate criterion-referenced statistics. Experts in statistical analysis report a deficit in this field. Consideration should be given to the rare-event and its influence on the results of analysis.

Additional Findings

8. Attempts should be made to determine error variations and their influence on learning. What is the frequency of false positive and false negative errors? Is there any difference in methods of correction for these errors? Do the error types differ according to the experience of the operator or for some other reason?
9. Further investigation should be made into the professional backgrounds of faculty to determine if there is a significant relationship between calculus detection skill level and background.

Closing Statement

Extensive research is needed on clinical evaluation in order to substantiate conclusions relating to validity. This study suggests that training may improve calculus detection skills of the dental hygiene instructors who participated. However, the conclusions are tenuous and the results yielded many unanswered questions.

Results of research on clinical evaluation have broad applications. In this study, at least three major areas could be effected by the results of the findings--both dental hygiene and dental formal educational institutions, dental licensing procedures, and continuing education programs. The models and training sequence could be used to improve calculus detection skills of dental hygiene and dental students, standardize licensing examiners, and provide continuing education to practitioners.

The results and the potential applications of research on validity in clinical evaluation necessitate further investigation.

APPENDIX A

ADMINISTRATIVE AND PARTICIPANT INSTRUCTIONS FOR CALCULUS DETECTION MODULE

Administrative Instructions for Calculus Detection Module

Introduction

Accurate calculus detection is a necessary skill for all dental hygiene instructors. This instructional module is designed to identify and improve the calculus detection for clinical instructors using dentition models.

Please read all of this packet first. Make sure everything is clear before starting. Call the program director if you have questions.

The module is designed to be primarily self-instructional with a pretest and a posttest. Supplies and equipment are minimal--dentition model, bench or chair mount, explorer, liquid soap, and recording forms.

As the module administrator you are asked to give each participant a pretest and a posttest. You are asked *not* to score them; just return the recording forms when they are completed by the participants. (Return envelopes enclosed.)

Instructions

Instructions for the pretest are as follows:

1. Introduce the concept of the module to the faculty by asking them to read the *Participant Instruction* directions (enclosed). This may be done as a group or individually.

2. Inform the participants that they will be taking a pretest and a posttest which consist of exploring and recording calculus on two dentition models.
3. Establish a place to administer the pretest and a series of convenient appointment times. Read the section entitled Suggestions for Testing Site to aid you in this task. As the administrator, you do not have to be present during the entire pretest, you are asked to help the participants get started, then you can allow them to continue alone at their own pace.
4. Ask the participants to turn in their recording forms to you when they are finished.
5. Repeat this same process for the posttest when the participants have completed their practice activities.

Suggestions for Testing Site

1. Identify a quiet area (one which can be isolated if possible) to stimulate clinical working positions.
2. Set up the two dentition models on mounts.
3. Have the following supplies set out:
   a. Recording form with participant number identified (See confidentiality section for explanation)
   b. Selection of different explorers
   c. Liquid soap

Pretest and Posttest Directions for Participants

1. Record your beginning and ending time in the space provided on the recording form.
2. Explore all of the dentition thoroughly. Approach this task as if you are evaluating a periodontally involved patient who has just received a complete scaling and rooting planing.
3. Mark on the dentition charts provided any calculus deposits you find; approximate their shape and location to the best of your ability.
4. Please do not attempt to visualize the deposits.
5. Apply liquid soap before beginning. This simulates saliva and it is important to maintain a film of soap during the entire exploring process. Squirt the soap onto the occlusals and rub it buccally and lingually.

Confidentiality

The results of the pretest and posttests will be analyzed statistically to identify initial and final levels of accuracy; however, please assure the participants that individual performance results will be confidential. As administrator, please do the following things to maintain confidentiality.

1. Assign each participant a number.
2. Identify the pretest and posttest recording forms by number only.

3. Retain the number and name list; do not send that to me.

After receiving the pretest recording forms, I will send you the group reliability scores. You should share these with your participants.

After the posttest, I will again send you the group scores. In addition, if any individuals would like their individual scores please request them by participant number. I will send them in individual envelopes to ensure privacy.

Please tell the participants about these confidentiality measures so that any feelings of threat are minimized. Emphasize improvement, not evaluation. The pretesting and posttesting is for my research; it is not designed to be punitive.

Practice Exercises

After the participants have completed the pretest, they are to practice their calculus detection skills. Reread the "Participant Instructions for Calculus Detection Models."

Participant Instructions for the Practice Exercises.

Make the dentition models and other supplies available to the participants to check in and out as they wish. Encourage them to practice as much as they can. All of the models you received may be used for practice except for the two identified as the testing models.

The testing models will be used for both pretest and posttesting. Please avoid telling the participants this to minimize memory interference.

Participant Instructions for Calculus Detection Module

Introduction

Calculus detection skills are of paramount importance to dental hygiene faculty for three purposes: patient care, student instruction and evaluation of student performance. This unit of instruction was developed for you, as an aid for refining your detection skills.

The module is a self-instructional package allowing for self-pacing and individualization; the author (a dental hygiene instructor) recognizes the time constraints and varying needs of busy faculty. Please adapt the activities of the module to fit your schedule.

The design format used for this unit is an eclectic one but the primary theory used was derived from the research on conceptual learning. Recent and extensive studies have identified specific patterns of learning sequences which facilitate learning; these patterns were used to develop the activities of the module.

Activities

Practice, practice, practice--is the main activity; the module has little written material. The goal is to provide you with an opportunity to systematically explore for calculus deposits on a model which simulates a "dirty" mouth as closely as possible. In addition you will have immediate feedback on your performance.

The package includes models and recording forms. You may use any explorer you feel comfortable using. Please use a bench or chair mount if you can, to simulate clinical positions.

Begin the module by taking a pretest with the module administrator. After completing this, begin your practice program. Directions for practice are simple:

1. Set up the models on the mount to approximate actual clinical positions.
2. Select the explorer(s) you wish to use.
3. Apply any liquid soap to simulate saliva; squirt it onto the teeth and rub it around. Reapply whenever you want to maintain a film throughout your practice. It is not necessary to apply soap under the gingiva because it will ooze under as you explore.
4. Explore the teeth in the following order.
    a. Quadrant 3, begin with 31 and work posteriorly
    b. Quadrant 4, begin with 41 and work posteriorly
    c. Quadrant 2, begin with 21 and work posteriorly
    d. Quadrant 1, begin with 11 and work posteriorly

Please follow this order because the conceptual patterns mentioned previously are built into the order indicated above

5. Record the presence of calculus by marking the appropriate circle on the recording form.
6. Your method of exploring each tooth is not critical, (i.e. fulcrum, grasp, stroke); however, try not to push or pull with excessive force. The calculus can be removed.
7. After completing one quadrant, check yourself by pulling up the gingiva and looking for calculus. Go back and re-explore if you missed something.
8. Continue through all of the quadrants on one model. When you complete them all, trade in your model for another one.
9. Be sure to give yourself a break of at least one hour between models. You do not have to complete all of the practice at one session.
10. Work through at least four different models. If you feel confident and your scores are consistently high--no more than three surface errors (incorrect responses per tooth surface) per model, ask to take the posttest. You may practice on as many models as there are models available, and you may repeat practice on models because memory retention is low from time to time.

You are encouraged to set up practice times convenient for yourself. Practice anywhere or anytime you decide is good for your schedule. The only restriction on practice time is imposed due to the limitations of this research. I ask that all participants complete the posttest with the administrator within three to four weeks after starting.

RECORDING FORM

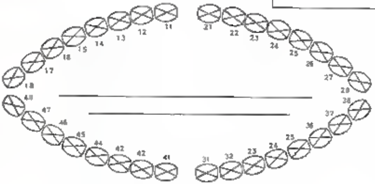
[illegible] ____________

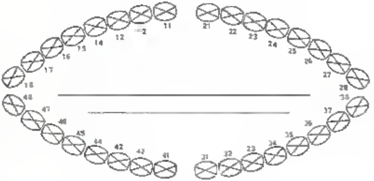
[illegible] ____________

For Pretesting and
Posttesting only.

Starting
Time ____________

Completion
Time ____________


11
12
13
14
15
16
17
18
21
22
23
24
25
26
27
28
48
47
46
45
44
42
42
41
31
32
23
24
25
36
37
38


11
12
12
14
15
16
17
18
21
22
23
24
25
26
27
28
48
47
46
45
44
42
42
41
31
32
23
34
35
36
37
38

APPENDIX B

SURVEY

May 5, 1983

Dear Project Administrator,

Enclosed you will find two questionnaires which are the final phase of the clinical evaluation research in which you have been participating.

Please have each of the participants complete a questionnaire. I have enclosed copies of the original instructions you received in case you or a faculty would like to refer to them.

The second questionnaire is for you, the project administrator.

Encourage participants to be as honest and wordy as they want to be. I appreciate any and all feedback.

Please return these to me by May 22.

Thank you for your cooperation. I will forward the results as soon as possible.

Sincerely yours,

Judy Skelton

Enclosures

Dear Faculty,

This questionnaire is the final phase of the clinical evaluation research in which you have been participating. The training portion of the project is of vital importance to the outcome; therefore, please answer the following questions as accurately and thoroughly as possible.

Yes / No Did you do any practice exercises?

If No, please explain why.

_____ Models inconveniently set up.

_____ No time during my teaching schedule.

_____ Did not feel the practice exercises would improve my detection skills.

_____ I had no real interest in the project.

_____ Practice exercises were too time consuming.

Other (specify)______________________________

____

____

If Yes, please describe the practicing you did.

_____ Estimate the time you spent doing practice exercises (in hours).

_____ What kind of explorer did you use?

Yes / No Did you use soap?

Yes / No Did you follow the order indicated in the instructions (begin with quadrant, 3, then, 4, 2, 1)?

Yes / No Did you check yourself after each quadrant?

1 2 3 4 ____ How many different models did you practice on?

Please check any or none of the following statements that are pertinent.

I feel the project would have improved my evaluation skills if there was (were)

_____ more bone loss on the models

_____ less bone loss on the models

_____ more calculus on the models

_____ less calculus on the models

_____ no gingiva for some of the practice time

_____ an initial introductory seminar

_____ specific supervised practice times

_____ review on calculus detection technique

Please add any other suggestions and comments you can make toward improvement of this project.

____

____

____

____

____

THANK YOU! THANK YOU! THANK YOU!

APPENDIX C

INSTRUCTIONS FOR DENTITION MODEL CONSTRUCTION

I. Materials and Equipment

A. Teeth - Set of (28) permanent teeth ordered from Unitec Corporation. Features of the teeth include coronal and root anatomy, life-like sizing, and color differentiation of roots and crown.

B. Latex - A liquid latex compound ordered from Cementex Company, Inc. The material is originally white and viscous; when exposed to air it solidifies into a pliable, rubbery material.

C. Rope Wax - Dental laboratory soft rope wax.

D. Instruments - #7 wax spatula, cleoid/discoid carver.

E. Lab Tile - Flat surface for model construction.

F. Alcohol Torch - Lab-type alcohol torch with blower which uses denatured alcohol.

G. Disposable Syringes - Plastic aspirating syringes.

H. Dental Plaster - Gypsum material of medium hardness.

I. Lathe - Bench-type model trimmer.

J. Epoxy Cement - Epoxy adhesive material which has color and which can be applied in small amounts.

II. Procedures for model construction.

A. Obtain an old set of study models. Trim off the teeth (using a model trimmer) maintaining the palatal and lingual sections.

B. Arrange the teeth in a semicircle on your desk in the correct order for placement in the arch. Be careful and review your dental anatomy.

C. Place a 1 cm thickness of wax around the periphery of the lingual and palatal arches.

D. Stick the teeth into the wax approximating an ideal arch--remembering to include correct contacts.

E. Using the alcohol torch and spatula, smooth the wax, duplicating the alveolar process. Make sure the wax is at least 4 mm from the cementoenamel junction (CEJ) and that the interproximal crests are approximately formed. The coronal portion of the alveolar process should be thin and contoured to the tooth.

F. Create malalignment.

  a. extract a tooth leading to tilting

  b. super eruption of a molar from a missing opposing tooth

  c. Crowding in anteriors with overlapping

  d. open contacts

  e. rotated and/or teeth in bucco or linguoversion

    f. roots in close proximity
    g. any other defect--seek prior approval

G. Carve out at least 5 bony defects.

H. Using a pencil, mark the height of contour of the bone throughout the arch. (This is to show you where to place the ligament later.)

I. Flow latex over the entire arch except for the root apices. Apply the latex in layers letting each layer dry between applications. Make sure the latex flows interproximally reproducing papillary contours.

J Mix up some dental stone, make a patty and invert the model into the stone (latex, teeth and wax).

K. After the stone sets, dig out the wax and teeth; remove the latex mold and clean all the wax off the teeth and mold. (Use hot water.)

L. Cut the interproximals off the mold.

M. Place orthodontic rubber bands on the teeth apical to the line marked previously.

N. Replace the teeth into the latex model and the latex into the stone holding mold

O. Flow plaster into the mold around all the roots to reproduce osseous contours.

P After the plaster sets, remove the arch from the mold and touch up the contours. Be careful not to dislodge the periodontal ligament (rubber bands), but make sure one surface is exposed.

Q. Create the gingiva by injecting latex first into the interproximals. Be careful not to create too much (hyperplastic) tissue. Apply in layers, letting the previous one dry.

R. Create a base for each arch which will fit into a bench-mounting articulator frame.

S. Apply calculus in small amounts. Mix up the epoxy cement provided and apply with a toothpick.

REFERENCES

Abou-Rass, M. (1973) A clinical evaluation instrument in endodontics. *Journal of Dental Education*, *37*(9), 22-36.

Ary, D., Jacobs, L.C., and Razavieh, A. (1979) *Introduction to research in education*. New York: Holt, Rinehart, and Winston.

Ausubel, D.P. (1968) *Educational psychology: A cognitive view*. New York: Holt, Rinehart, and Winston.

Baratti, N.J.C. (1974) *The homogeneity of grading practices of a dental hygiene clinical faculty*. Unpublished master's thesis. University of Michigan, Ann Arbor, Michigan.

Biller, J.R., and Kerber, P.E. (1980) Reliability of scaling error detection. *Journal of Dental Education*, *44*(4), 205-210.

Bilodeau, I. (1966) Information feedback. In E. A. Bilodeau (Ed.) *Acquisition of skill*. New York: Academic Press.

Bloom, B. S. (1956) *Taxonomy of educational objectives*. New York: David McKay Company.

Bruner, J. S., Goodnow, J.J., and Austin, G.A. (1956) *A study of thinking*. New York: Wiley.

Casteel, J.D., Button, C.B., Corbett, W.T., Gregory, J.W., LaCava, H. J., and Stahl, R.J. (1974) *Using the concept of "competition" to develop problem-solving skills*. Florida State Department of Research and Development Report, University of Florida, Gainesville, Florida.

Centra, J. A. (1976, November) *Faculty development practices in United States colleges and universities*. Princeton, NJ: Educational Testing Service.

Cohen, J. A. (1960) A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37-46.

Davies, R.H., Alexander, L.T., and Yelon, S.L. (1974) Learning system design. New York: McGraw-Hill Book Company.

Fitzpatrick, R., and Morrison, E.J. (1971) Performance and product evaluation. In R.L. Thorndike (Ed.) Educational Measurement. (pp. 197-204) Washington D.C.: American Council on Education.

Fleishman, E.A. (1962) The description and prediction of perceptual motor skill learning. In F. Glaser (Ed.) Training, Research and Education. (pp. 137-175). Pittsburgh: University of Pittsburgh Press.

Francis, J.B. (1975) How do we get there from here? Program design for faculty development. Journal of Higher Education, 46(6), 719-732.

Fuller, J.L. (1972) The effects of training and criterion models on interjudge reliability. Journal of Dental Education, 36(4), 19-22.

Gaff, J.G. (1975) Toward faculty renewal. San Francisco: Jossey-Bass, Publishers.

Gagne, R.M. (1977) The conditions of learning (3rd ed.) New York: Holt, Rinehart and Winston.

Gagne, R.M., and Briggs, L.J. (1974) Principles of instructional design. New York. Holt, Rinehart and Winston.

Gaines, W.G., Bruggers, H., and Rasmussen, R.H. (1974) Reliability of ratings in preclinical fixed prosthodontics: Effect of objective scaling. Journal of Dental Education, 38(12), 672-675.

Goepferd, S.J., and Kerber, P.E. (1980) A comparison of two methods for evaluating primary class two cavity preparations. Journal of Dental Education, 44(9), 537-542.

Gronlund, N.E. (1978) Starting objectives for classroom instruction (2nd ed.) New York: Macmillan Publishing.

Hall, D.C., Goldstein, M.K., and Steen, G.R. (1977) Progress in Manual Breast Examination. Cancer, 40(1), 264-370.

Hammons, J., Smith, T.E., and Watts, G. (1979) *Staff development in the community college: A handbook.* (Topical paper No. 66) Los Angeles: ERIC Clearing House for Junior Colleges (ERIC Documentation Reproduction Service No. 154 987).

Hartman, D.P. (1977) Considerations in the choice of interobserver reliability estimates. *Journal of Applied Behavioral Analysis*, *10*(1), 103-116.

Hinkelman, K.W., and Lang, N.K. (1973) Method for decreasing subjective evaluation in preclinical restorative dentistry. *Journal of Dental Education*, *37*(9), 13-19.

Houpt, M.I., and Kress, G. (1973) Accuracy of measurement of clinical performance in dentistry. *Journal of Dental Education*, *37*(7), 34-36.

Hunter, H.G., Salkin, F.M., Love, R., and Hildebrand, C.N. (1975) Deriving clinical performance standards. *Journal of Dental Education*, *39*(10), 651-657

Irion, A. (1966) A brief history of research on the acquisition of skill. In E.A. Bilodeau (Ed.), *Acquisition of skills* (pp. 14-36). New York: Academic Press.

Joyce, B.R., and Weil, M. (1980) *Models of teaching* (2nd ed.). Englewood Cliffs, New Jersey: Prentice-Hall.

Keeton, M.T. (1976) *Experimental learning: Rationale, characteristics and assessment.* San Francisco: Jossey-Bass.

Klausmeier, H.S., and Feldman, K.V. (1975) Effects of a definition and a varying number of examples and non-examples on concept attainment. *Journal of Educational Psychology*, *67*(2), 174-179.

Knowles, M.S. (1979) *The adult learner* (2nd ed.). Houston: Gulf Publishing Co.

Levine, A. (1978) *Handbook on undergraduate curriculum.* San Francisco: Jossey-Bass.

MacKenzie, R. (1973) Defining clinical competence in terms of quality, quantity and need for performance criteria. *Journal of Dental Education*, *37*(9), 37-44.

MacKenzie, R. (1974) Factors essential to evaluation of clinical performance. *Journal of Dental Education*, *38*(4), 214-223.

Mackworth, J. (1964) The effect of true and false knowledge of results on the detectability of signals in a vigilance task. *Canadian Journal of Psychology*, *18*(2), 106-117.

Martorella, P. (1972) *Concept learning: Designs for instruction*. New York: Intext Educational Publishers.

McKinnan, J.W. (1976) The college student and formal operations. In T. Renner (Ed.), *Research training, teaching and learning with the Piaget model* (pp. 110-129). Norman, Oklahoma: University of Oklahoma Press.

Mehrens, W.A. and Lehman, I.J. (1973) *Measurement and evaluation in education and psychology*. New York: Holt, Rinehart and Winston.

Merrill, M.D., and Tennyson, R.D. (1977) *Teaching concepts: an instructional design guide*. Englewood Cliffs, New Jersey: Educational Technology Publication.

Mertens, R.P. (1971) The development and evaluation of a self-instructional course in braille-micrometer reading. (Doctoral Dissertation, Pennsylvania State University, 1971). *Dissertation Abstracts International*, *32*, 2439A

Morgan, M.K., and Irby, D.M. (1978) *Evaluating clinical competence in the health professions*. St. Louis: C.V. Mosby Co.

Natkin, E., and Guild, R.E. (1967) Evaluation of preclinical laboratory performance: a systematic study. *Journal of Dental Education*, *31*(2), 152-161

Norton, R.E. (1967) Using programmed instruction with and without self-instructional practice to teach psychomotor skills (Doctoral dissertation, Cornell University, 1966). *Dissertation Abstracts International*, *28*, 2603A.

O'Banion, T.D. (Ed.) (1977) Developing staff potential. In *New directions for community colleges*, Publication 19. San Francisco: Jossey-Bass

O'Connor, P., and Lorey, R.E. (1978) Improving interrater agreement in evaluation in dentistry by the use of comparison stimuli. *Journal of Dental Education*, *42*(4), 174-177.

Ormes, W.M. (1973) Evaluation of the student in clinical periodontics. *Journal of Dental Education*, *37*(10), 42-43.

Parker J., and Fleishman, E. (1961) Use of analytical information concerning task requirements to increase the effectiveness of skill training. *Journal of Applied Psychology*, *45*(5), 295-302.

Patridge, M.J., and Mast, T.A. (1978) Dental clinical evaluation: a review of the research. *Journal of Dental Education*, *42*(6), 300-305

Pattison, G.L., and Pattison, A.M. (1978) *Periodontal Instrumentation*. Reston, Virginia: Reston Publishing.

Phillips, S.R. (1976, May) What, then, is faculty development? *Association of the Departments of English Bulletin*, p. 11-17

Piaget, J., and Inhelder, B. (1969) *The psychology of the child*. New York: Harper.

Rapport, R., and Rapport, R.N. (1975) *Leisure and the family life cycle*. London: Rutledge and Kegan-Paul.

Razzoog, M., Shotwill, J.L., and Lang, B.R. (1978) Performance criteria in preclinical denture prosthodontics. *Journal of Prosthetic Dentistry*, *40*(5), 576-578.

Schluger, S., Yuodelis, R., and Page, R.C. (1977) *Periodontal disease*. Philadelphia: Lea & Febiger

Seigel, S. (1956) *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill Book Company.

Simpson, E.J. (1972) The classification of educational objectives in the psychomotor domain. *The Psychomotor Domain* (Vol. 3). Washington: Gryphon House.

Smith, A.B. (1976) *Faculty development and evaluation in higher education*. Higher education research report (ERIC Report No. 8). Washington: American Association for Higher Education. (ERIC Document Reproduction Service No. 132 891).

Smith, K., Wargo, L., Jones, R., and Smith, W. (1963) Delayed and space displaced sensory feedback and learning. *Perceptual Motor Skills*, *16*, 781-796.

Spitzer, D.R. (1975) What is concept? *Educational Technology*, *15*(7), 36-39.

Stahl, R.J. (1979, April) *Validating a modified Gagnean concept acquisition model*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, California.

Stahl, R.J., Button, C.B., and Corbett, W.T. (1975) *A model to create instructional activities intended to develop student concept learning and problem-solving skills: A modification of Gagne's "Conditions of Learning" for social studies instruction*. Paper presented at the Annual Meeting of the American Education Research Association, Washington, D.C. (ERIC microfiche no. ED 106174).

Tennyson, R.D., and Park, O. (1980) Teaching concepts: A review of instructional design research literature. *Review of Educational Research*, *50*(1), 55-70.

Troy, C.L. (1971) *Periodontal instrumentation: Tactile discrimination*. Unpublished master's thesis, University of Michigan, Ann Arbor.

Wiener, E. (1963) Knowledge of results and signal rate in monitoring: A transfer of training approach. *Journal of Applied Science*, *47*(3), 214-222.

Wilkins, E.M. (1976) *Clinical practice of the dental hygienist*. Philadelphia: Lea and Febiger.

Williams, R.E (1981) Self-instruction in dental education: 1960-1980. *Journal of Dental Education*, *45*(5), 290-299.

Wood, R.K., and Neil, J.T. (1978) Experimental Learning An alternative approach to staff development. *Texas Technical Journal of Education*, *5*(2), 113-122.

Wood, F.H., and Thompson, S.R. (1980) Guidelines for better staff development. *Educational Leadership*, *37*(5), 374-378.

Woodall, I.R., Dafoe, B.R., Young, N.S., Weed-Fonner, L., and Yankell, S.L. (1980) *Comprehensive dental hygiene care*. St. Louis: C.V. Mosby.

BIOGRAPHICAL SKETCH

Judith Skelton was born in Charleston, South Carolina, on June 7, 1950. She received her Bachelor of Science degree in dental hygiene from the University of Louisville, in 1974. After working for one year in a dental research center in Louisville, she entered the Peace Corps. Upon returning to the United States she obtained employment as a dental hygienist in a private dental office and later at the University of Florida, College of Dentistry.

After one year, Ms. Skelton began teaching in the Dental Auxiliary Programs at Santa Fe Community College. She concurrently entered the University of Florida, College of Education and in 1977 obtained a Master of Education in adult, vocational and technical education with an emphasis on health related education.

In 1978, Ms. Skelton began a doctoral degree in curriculum and instruction with an emphasis in dental education During her graduate study she continued teaching in the dental auxiliary programs at Santa Fe Community College.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Albert B. Smith
Albert B. Smith, Chairperson
Professor of Instructional
Leadership and Support

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Richard S. MacKenzie
Richard S. MacKenzie
Professor of Instructional
Leadership and Support

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy

Wallace L Mealiea
Wallace L. Mealiea
Professor of Clinical
Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Margaret K. Morgan
Margaret K Morgan
Professor of Instructional
Leadership and Support

This dissertation was submitted to the Graduate Faculty of the Division of Curriculum and Instruction in the College of Education and to the Graduate Council, and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy

December, 1983

Dean for Graduate Studies and
Research